Statewide Evaluation: Replicating Science-based Substance Abuse Prevention Programs[1]

November 2005

Rhode Island Department of Mental Health Retardation and Hospitals

Division of Behavioral Healthcare Services

Prepared by the[2]

Community Research and Services Team
at the University of Rhode Island

[1] The RI Department of Mental Health, Retardation, and Hospitals with funding from the Substance Abuse and Mental Health Services Administration (SAMHSA) Substance Abuse Prevention and Treatment (SAPT) Block Grant provided support for the collection of the data on which this report is based via a contract for Statewide Evaluation of the Replication of Science-based Programs.

[2] Community Research and Services Team, Department of Psychology, University of Rhode Island, Providence Campus, Room 236, 80 Washington Street, Providence, RI 02903-1803. (401) 277-5318

EXECUTIVE SUMMARY

This four-year project investigated the effectiveness of three well-researched "model" programs for the prevention of alcohol, tobacco, and other drug abuse. Two of the programs were designed to be delivered in middle school classrooms over a number of weekly sessions: Life Skills Training (LST) and Project Northland (PN). The third program was designed to be delivered to parents/guardians and their middle school child: Iowa Strengthening Families Program (ISFP). With funding from the Division of Behavioral Healthcare Services, RI Department of Mental Health, Retardation and Hospitals, five community-based organizations mounted these interventions in eight middle schools spread across seven school districts in RI. At each site one of the classroom curricula was combined with the family-oriented program for implementation over a three-year period with one cohort of middle school students.

Did the programs reach the intended target population?

- 1,326 students had received three years of exposure to one of the classroom prevention curricula at the time we conducted outcome analyses. There is enough statistical power to detect even modest effects for both of the classroom curricula.

- The Iowa Strengthening Families Program had 98 youth and 106 parents for whom we could analyze results, representing only 7.4% of the classroom treatment sample

.

- Eighth graders in our outcome analyses were approximately 50% female, and varied across the schools, with program-level subgroups ranging from 67.5% to 83.6% white, 55.3% to 79.7% ineligible for subsidized lunch, 71.5% to 79.7% in two-parent families, and 57.1% to 63.5% with grades mostly B or better.

- The non-random way in which schools chose the curricula makes any competitive comparisons of the two classroom curricula suspect; for example, 75.8% of the Project Northland treatment group paid full price for their lunches, whereas only 55.3% of the Life Skills Training treatment group did so.

- Students in the ISFP program were more likely to be white and less likely to be in two-parent families than the youth treatment cohort as a whole.

Were the programs delivered with fidelity?

- All three programs were delivered with relatively good fidelity: ISFP (94.9%), PN (88.4%), and LST (80.1%).

Were our outcome measures effective?

- We measured performance on Intermediate Objectives most likely to show direct effects of the programs (e.g. perceived peer disapproval for use, attitudes toward drug use, drug refusal skills, peer normative beliefs, intentions to use drugs, family attachment, and parental monitoring. Psychometric properties were reasonable for our sample.

- We measured Outcome Objectives required by RIMHRH (30-day prevalence of substance use, age of initiation of use, and problem drinking). We particularly examined alcohol use as an appropriate indicator for this age group.

Did the programs achieve intermediate objectives by eighth grade?

- When we compared our 8th grade treatment group to untreated 8th graders from the same schools, there were highly significant positive differences; the largest differences were for Favorable Attitudes Toward Drug Use and Drug Use Intentions (both lower for our treatment group).

- All four intermediate objectives specific to Project Northland showed modest but significant effects in the right direction, including effects on youth perception of parents (quality of parent communication and rule enforcement for ATOD use).

- For Life Skills Training, three of five intermediate outcome measures showed modest positive effects (higher drug refusal skills, reduced pro-drug attitudes, and lower perceived peer norms).

- For ISFP, parents indicated significant improvement over time on all five outcomes but only one intermediate outcome showed significant positive change for both parents and youth in ISFP: there was a significant increase in "Limit Setting and Monitoring." This is worth celebrating, as it is an important protective factor.

Did the programs achieve effects on substance use outcomes?

- For 30-day prevalence of alcohol use, probably the most widely chosen indicator for studies with this age group, both programs produced substantial effects (45% lower alcohol use than for the comparison group) that were highly significant and did not differ between the two classroom curricula.

- SALT data confirmed this finding.

- For initiation of alcohol use during the three years of the programs, there was a significant effect when both programs were combined but this was due to the substantial effect of Project Northland (42% lower initiation than the comparison group) and did not show up for the Life Skills Training intervention.

- Both programs had a significant effect on problem drinking. Eighth grade youth in the comparison group had a 10.9% rate of problem drinking, while the eighth grade youth in the treatment group had a 6.5% rate. This represents a 40% lower rate.

- The Iowa Strengthening Families Program did not have a significant added effect. However, there are important qualifications for this conclusion, including the small sample size and the weak self-report measure we had of participation in ISFP for these analyses.

Final conclusions

- If the goal is decreasing initiation of use (any experimentation with alcohol) PN is a more promising choice, especially for non-white and lower S.E.S. students.

- If reducing 30-day prevalence (regular use of alcohol over time) is the goal, both programs did very well and LST was especially effective for white and higher S.E.S. students.

- For reducing problem drinking (three or more drinks at one time in the past two weeks) both programs did very well.

Acknowledgements

The CRST thanks superintendents and principals in participating school districts and schools for allowing us to evaluate the programs offered in their schools. Their interest and support were essential to the project's success. We also want to thank the Community Substance Abuse Prevention Task Forces and Student Assistance Counselors who facilitated the survey administration process in some communities, as well as the teachers and other providers who gathered information from students and some of their families, who provided information on a daily basis in the form of fidelity checklists and who provided self-evaluation data regarding their training in the model program curriculum. We also thank the vendor agencies and their dedicated staff who diligently fulfilled active roles beyond the scope of this project and those who worked behind the scenes. Continued concern and cooperation at the local level was essential to the success of the Science-based Demonstration Project. We hope the outcome of this project will justify current and future participation.

We especially thank the thousands of students who responded voluntarily to participate in taking the youth survey. Their willingness to provide us with three years of information about themselves is greatly appreciated. We thank parents and their youth who participated in the family-based program for providing information as well.

The Rhode Island Department of Mental Health, Retardation and Hospitals, Division of Behavioral Healthcare Services, provided support for this project, with funding from the U.S Center for Substance Abuse Prevention.

Table of Contents

Background for this project

Overview

Over the past two decades there has been a move toward an increased emphasis on accountability of human service programs in the United States. During this time public and private sectors funded research programs that have been developing a science of prevention. The Division of Behavioral Healthcare Services (DBHS), Rhode Island Department of Mental Health, Retardation, and Hospitals funded demonstration projects that combined replications of three "science-based" model programs for the prevention of alcohol, tobacco, marijuana, and other drug abuse for youth ages 10-14 and a portion of their families. The Division awarded contracts to five community-based organizations in documented partnerships with one or more middle schools (totaling eight middle schools). The Community Research and Services Team (CRST), based in the Psychology Department at the University of Rhode Island, served as the statewide evaluators.

Scope of Work Performed

As the statewide evaluators, the CRST provided direction and guidance to the vendor agencies in the management of their evaluations, prepared and supplied pre-post instruments used by vendors to measure intermediate and ultimate outcomes for the programs and a tool for monitoring the fidelity of each of the programs, provided technical assistance in program evaluation to the vendors through a designated liaison, coded and analyzed the data collected by the local vendors, provided feedback to the vendors in the form of data summaries, and compiled outcome data across the demonstration projects to prepare this final report.

Description of the Prevention Programs Used

The funded vendors chose to implement one of two youth-oriented prevention curricula for their projects. The Life Skills Training (LST) Program is a school-based tobacco, alcohol, and drug abuse prevention program for adolescents in grade six with booster sessions continuing in grades seven and eight. Project Northland (PN) is a community-wide program (with peer-led school-based curriculum, parent involvement, and a community-wide policy change component) sequentially designed for presentation to students in grades six through eight. Each of the funded vendors elected to implement the Iowa Strengthening Families Program (ISFP) as a family-based prevention program that focuses on improving parent-child relationships through changing family dynamics and helping families work together as a unit. A more detailed description of each program can be found in Appendix A.

Objectives and Logic Models

DBHS specified the following outcome objectives for youth reached by the funded demonstration projects:

Objective 1: The percent of treated eighth graders who report initiation of tobacco and alcohol use will be 10% lower than an untreated comparison group by Year 3 of the project;

Objective 2: The percent of treated eighth graders who report current use of tobacco, alcohol, marijuana and inhalants (30 day prevalence) will be 10% lower than an untreated comparison group by Year 3 of the project;

Objective 3: The percent of treated eighth graders who report current problem drinking will be 10% lower than an untreated comparison group by Year 3 of the project.

In addition to the DBHS-required outcome objectives, the selected programs have documented effects on intermediate objectives. The CRST, in conjunction with the local agencies, identified relevant intermediate objectives and selected promising measures for these objectives, considering the core measures recommended by CSAP and the locally available sources for comparison, such as the Youth Tobacco Survey and the SALT. Some measures were shared between LST and PN and others were specific to the components of one of the programs. Figures 1, 2 and 3 present program-specific logic models that provide descriptions of these objectives.

Statewide Evaluation Design

The statewide outcome evaluation design made use of self-report questionnaires for youth in the LST and PN programs. Youth and parents who participated in the family based program, ISFP, were also measured using pre and post survey questionnaires. Copies of the instruments used can be viewed in Appendix B. Surveys were administered before and after each multi-session year of the LST and PN curriculum, tracking the same cohort from sixth to eighth grade. A unique identifier code known only to the participant was used to track individual respondents over time (see Appendix C for the matching protocol, and a paper discussing this procedure). This anonymous technique of matching data allowed for repeated measures analyses and examination of attrition effects. A sample consisting of all eighth graders from the same schools in year one served as a comparison group. These students responded to the survey, but were not exposed to the curriculum. Data from the comparison group were collected in spring 2003. An investigation of 30-day prevalence for alcohol use by month for the comparison group indicated periods of increased reports of use and non-use. In order to provide stable measurements for comparison, the treatment groups responded to the same instruments at approximately the same time of year as their comparison group for their final post-test when they reached the eighth grade in 2005. The effects of participation in ISFP are examined as an enhancement to the classroom curricula. In addition to changes over time for students, comparisons are made between the eighth grade cohort and eighth grade comparison group, between LST and PN, between participants with and without exposure to the ISFP, and results are also compared to state trends obtained from the SALT. Effects are analyzed controlling for demographic differences, and specific demographic factors are examined to investigate the effectiveness of programs for diverse groups.

Four of the five participating vendors elected to use passive consent. Prior to implementation, letters written by each vendor were mailed to the parents of all sixth and eighth grade students in each of the supporting schools. Parents were directed to return the letters if they did not want their children to participate in the survey. Participation in the survey was voluntary. Even if parents did not return the letters indicating their refusal, students were given the opportunity to assent to or decline completing the survey. The agency that elected to use active consent followed the same protocol as the others; however, the parents of each student in their school needed to return the letter indicating approval for their children to respond to the survey. Youth of parents who did not return the letter did not complete the surveys. Parents and youth participating in the ISFP gave oral agreement to participate in the survey at the first meeting.

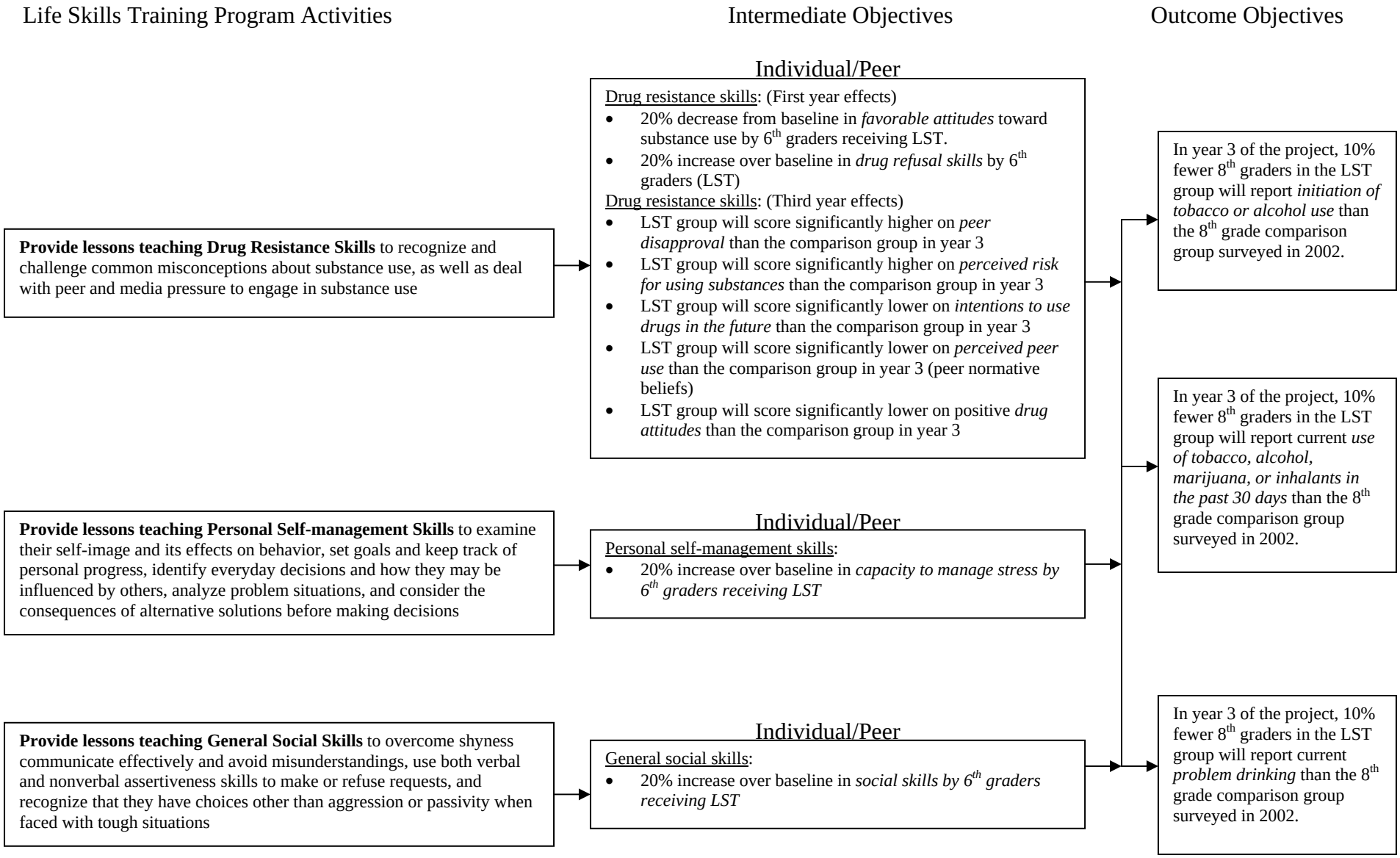

Figure 1. Life Skills Training Logic Model

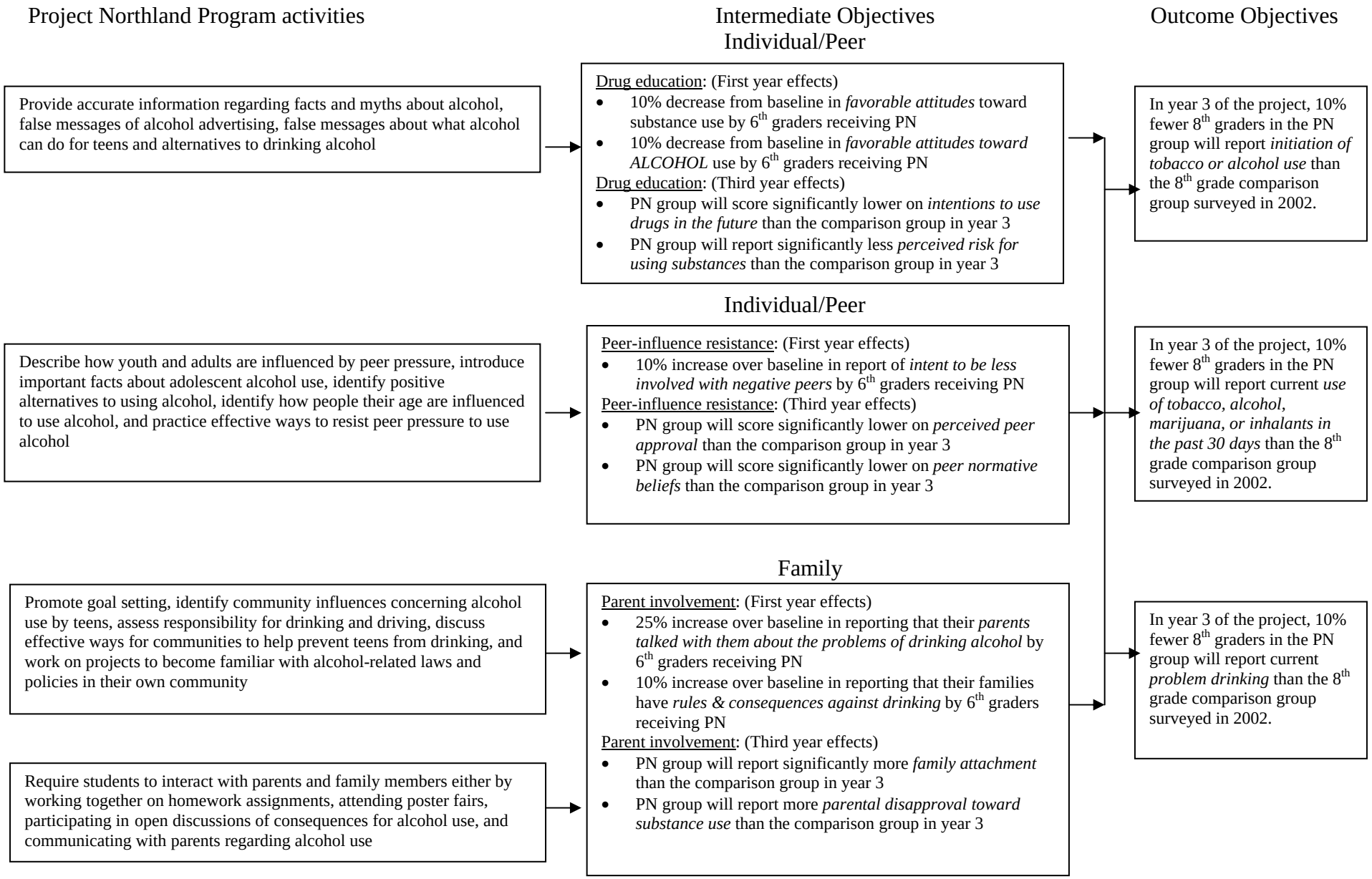

Figure 2. Project Northland Logic Model

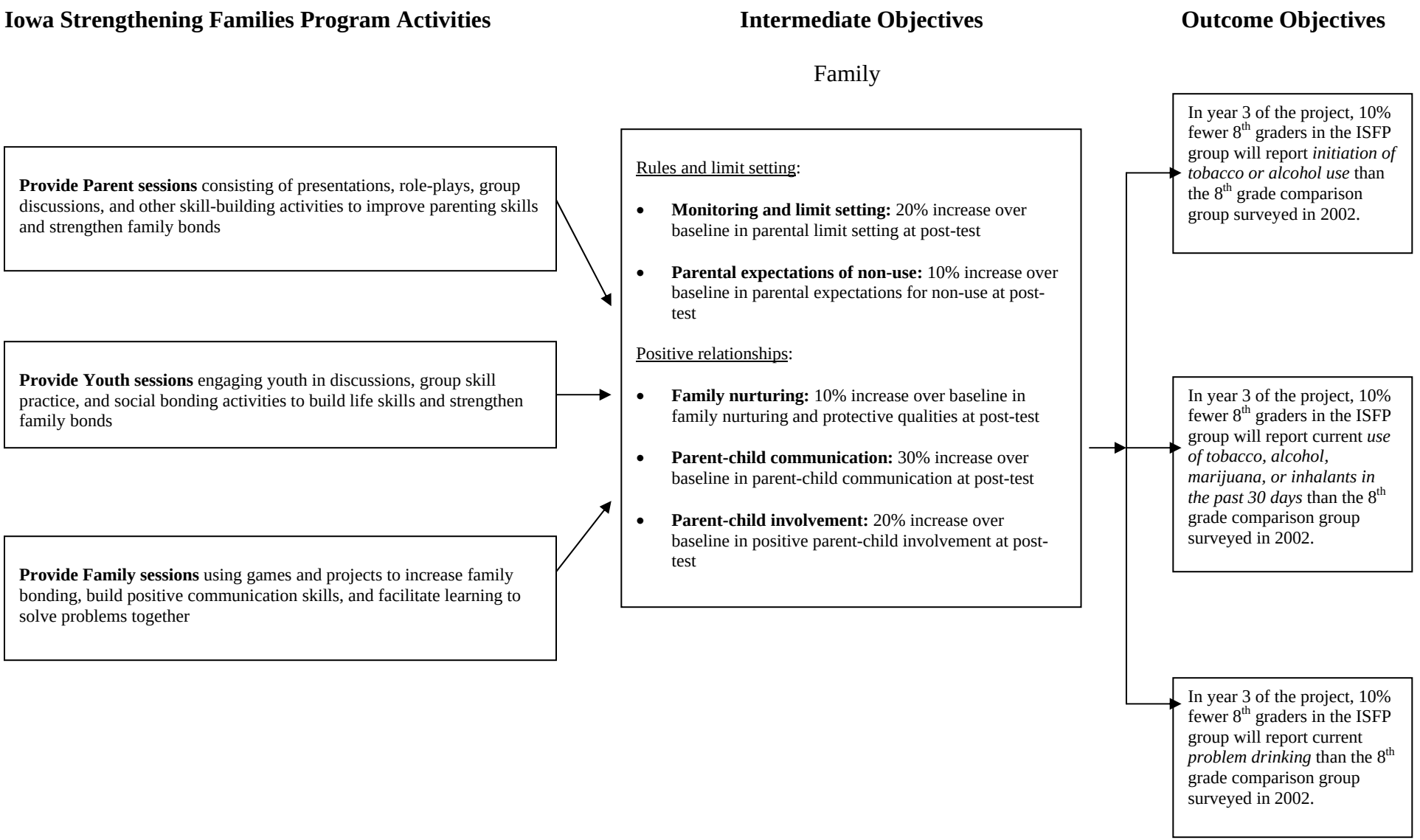

Figure 3. Iowa Strengthen Families Program Logic Model

Process Evaluation of the Program

In this section we will report on the characteristics of youth and parents who participated in the programs, the fidelity of treatment, our evolving process for coordinating the evaluation and the initial training of program staff in curriculum delivery.

Demographic characteristics of the samples used in analyses

PN and LST

Table 1 provides information on demographic characteristics, contrasting the eighth grade comparison group with the eighth grade treatment group (the entire sample receiving 3 years of the prevention curriculum). Youth in the treatment cohort who did not receive the program in its entirety or who left the school were not surveyed in the eighth grade. The comparison group is comprised of 1443 respondents, while the treatment group provided data from 1326 participants. Treatment and comparison group sample sizes vary among the schools receiving the program; on average the Project Northland school comparison samples are bigger than the treatment group samples, while the reverse is true for the Life Skills Training schools. The percent of female participants in each school differed slightly, but gender was represented similarly for the treatment and comparison groups. Age was quite similar across schools and groups, although students at Calcutt Middle School (CMS) in Central Falls were slightly older on average. The percent of white students was similar for comparison and treatment groups, though slightly higher for the comparison group. CMS was noticeably different from all of the other schools in the project in race composition (32% White in the comparison group and 21% in the treatment group) and eligibility for free or partial subsidy for school lunch (3.7% Full pay in the comparison group and 1.2% in the treatment group), and this brought the overall averages for these characteristics down in the Life Skills sample.

Table 2 provides additional information on demographic characteristics, contrasting the matched sample (sixth graders in our treatment group whom we were able to track across seventh and eighth grades) with the unmatched sixth graders (those from whom we collected data at Time 1 but were unable to track through the subsequent two years). Participants we could not match were consistently lower in percent White, full pay lunch, coming from two-parent families, and obtaining grades of 'B' or better. As in the eighth grade samples described above, CMS contributed to a discrepancy between the Life Skills students and the Project Northland students on these characteristics.

ISFP

Table 3 provides information on demographic characteristics of the parent/care providers and youth participating in the family based program. Only those parents (66.3% of the total parent sample) and youth (76% of the total youth sample) who were present at both the pre- and post-tests for the initial 7-session intervention are included. Over the three years, 204 participants (106 parent/care providers and 98 youth) provided pretest and posttest data that we could successfully match. About three quarters of the parents were female (mostly mothers but also a few grandmothers). Our intention was to include only participants in our cohort and therefore we began by collecting data from the sixth graders in 2003 (55% of the sample) and continued in the following year with data from seventh graders (30%) and from eighth graders (15%) in year three. Note that these are non-overlapping groups of students whose families participated in the initial 7-session intervention during different years. Roughly 85% of parents were white; 60% did not qualify for subsidized lunch; 90% spoke English at home; the median years of education extended two years past high school; and about half were in 2-parent families. However, there was noteworthy variation between the schools on some of these demographic factors. Parents in Curtis Corner Middle School in South Kingstown (n = 13) and Pier Middle School in Narragansett (n = 9) were all White, while none of the parents of children in Calcutt Middle School in Central Falls (n = 6)

described themselves as White. A Spanish version of the parent survey was created and used in Central Falls for posttest due to the high percentage of parents who had difficulty with the English version. Although a majority of families from most of the schools paid full lunch, all parents from Central Falls received at least partial subsidy for their children's school lunch, and at J.F. Deering in W. Warwick 56% qualified for subsidy. In the initial classroom survey conducted in sixth grade, 78% said they were White and 75% said they were in 2-parent families. Thus, the families who chose to participate in ISFP were somewhat more likely to be White and less likely to be in 2-parent families than the youth treatment cohort as a whole. Anecdotal reports at liaison meetings indicated that a few divorced parents attended the sessions together. Of the 98 youth who were matched in the ISFP sample, 43% were female and the majority (75%) participated in the sixth grade.

Table 1. Demographics for 8th graders (2002) & 8th graders (2005) in Project Northland & Life Skills Training Program Schools							
School	**Demographic Characteristics for Treatment and Comparison Cohorts**						
Project Northland	Number Complete (N)	Female (%)	Age (Average)	White (%)	Full Pay Lunch (%)	Two-Parent Family (%)	Mostly B or Better (%)
Broad Rock Comparison	175	45.7	13.6	86.9	90.1	80.5	75.6
Broad Rock Treatment	142	55.3	13.6	89.4	87.3	84.4	70.4
Park View Comparison	272	56.1	13.6	74.6	72.7	66.4	48.3
Park View Treatment	160	45.1	13.9	59.4	59.5	71.2	41.2
Curtis Corner Comparison	155	49.4	13.6	91.0	93.5	82.5	68.4
Curtis Corner Treatment	161	42.1	13.5	92.5	90.1	83.9	82.6
J F Deering Comparison	235	53.7	13.5	86.8	70.8	75.3	45.9
J F Deering Treatment	206	53.2	13.5	85.0	69.3	75.5	54.3
Total (weighted) Comparison	837	52.0	13.6	83.6	79.7	74.8	57.1
Total (weighted) Treatment	669	49.0	13.6	81.6	75.8	78.4	61.4
Total Combined	**1506**	**50.5**	**13.6**	**82.6**	**77.7**	**76.6**	**59.2**
Life Skills Training							
Calcutt Comparison	142	65.2	14.4	32.1	3.7	65.9	44.4
Calcutt Treatment	171	54.3	13.6	20.5	1.2	53.3	59.1
Pier School Comparison	139	47.5	13.5	96.4	88.3	78.4	63.2
Pier School Treatment	93	48.3	13.7	92.5	88.9	80.2	65.9
Portsmouth Comparison	205	48.0	13.6	97.0	92.4	80.0	74.9
Portsmouth Treatment	215	46.4	13.5	92.6	89.5	84.4	76.2
Thompson Comparison	120	53.3	13.8	68.6	49.6	60.2	57.6
Thompson Treatment	178	48.0	13.6	69.1	48.3	69.0	51.1
Total (weighted) Comparison	606	53.0	13.8	76.0	62.2	72.4	61.6
Total (weighted) Treatment	657	49.2	13.6	67.5	55.3	71.5	63.5
Total Combined	**1,263**	**51.1**	**13.7**	**71.7**	**58.7**	**72.0**	**62.6**
Project Total (weighted)	**2,769**	**50.8**	**13.6**	**77.6**	**69.1**	**74.4**	**60.6**

Table 2. Sixth graders in the treatment cohort receiving Project Northland & Life Skills Training Program							
School	**Demographic Characteristics for Matched and Unmatched Youth**						
Project Northland	Number Complete (N)	Female (%)	Age (Average)	White (%)	Full Pay Lunch (%)	Two-Parent Family (%)	Mostly B or Better (%)
Broad Rock T1 T6 matched	134	57.9	11.2	88.8	86.5	85.1	91.0
Broad Rock T1 T6 unmatched	33	45.5	11.2	84.8	78.8	69.7	63.6
Park View T1 T6 matched	90	46.5	11.4	81.8	70.8	85.1	66.3
Park View T1 T6 unmatched	116	42.2	11.5	74.6	67.3	66.1	61.4
Curtis Corner T1 T6 matched	143	39.2	11.3	91.6	93.7	88.8	87.3
Curtis Corner T1 T6 unmatched	27	59.3	11.3	88.9	74.1	77.8	81.5
J F Deering T1 T6 matched	169	53.9	11.3	89.2	67.7	79.9	74.1
J F Deering T1 T6 unmatched	100	43.0	11.4	81.6	43.9	64.6	59.6
Total (weighted) Matched	536	49.7	11.3	88.5	79.9	84.4	80.5
Total (weighted) Unmatched	276	44.6	11.4	79.8	60.9	67.1	63.0
Total Combined	**812**	**47.1**	**11.4**	**84.1**	**70.4**	**75.8**	**71.8**
Life Skills Training							
Calcutt T1 T6 matched	85	61.0	11.6	23.5	1.2	61.2	64.5
Calcutt T1 T6 unmatched	77	60.8	11.6	20.8	0.0	59.7	65.7
Pier School T1 T6 matched	91	49.4	11.0	92.3	86.8	81.3	92.3
Pier School T1 T6 unmatched	30	58.6	11.0	90.0	73.3	50.0	72.4
Portsmouth T1 T6 matched	184	48.9	11.0	95.7	89.3	84.8	87.4
Portsmouth T1 T6 unmatched	34	35.3	11.0	88.2	93.8	79.4	84.4
Thompson T1 T6 matched	125	47.5	11.4	67.2	50.8	68.8	59.0
Thompson T1 T6 unmatched	70	42.9	11.5	64.3	37.7	45.7	40.0
Total (weighted) Matched	485	50.8	11.2	75.1	63.5	75.9	77.0
Total (weighted) Unmatched	211	50.4	11.4	55.9	38.0	56.9	61.1
Total Combined	**696**	**50.6**	**11.3**	**65.5**	**50.8**	**66.4**	**69.1**
Project Total (weighted)	**1,508**	**49.2**	**11.33**	**78.0**	**65.3**	**74.7**	**73.5**

Time 1 (T1) pretest grade 6 Time 6 (T6) posttest grade 8

Table 3. Participants in the Iowa Strengthening Families Program

School	Demographic Characteristics for Matched Parent & Youth								
ISFP PARENT	**Number Complete (N)**	**Female (%)**	**White (%)**	**Full Pay Lunch (%)**	**English at Home (%)**	**Median Years Education (%)**	**Married, Spouse absent (%)**	**Married Spouse Present (%)**	**Unmarried (%)**
Broad Rock matched	17	52.9	94.1	70.6	94.1	16.0	0	81.3	18.8
Calcutt matched	6	100.0	0.0	0.0	16.7	12.0	16.7	0	83.4
Park View matched	29	92.9	89.7	55.2	96.6	13.0	20.7	31	48.2
Curtis Corner matched	13	66.7	100.0	84.6	100.0	17.5	0	76.9	23.1
Pier School matched	9	66.7	100.0	77.8	100.0	15.5	0	42.9	57.1
Portsmouth matched	6	66.7	83.3	83.3	83.3	16.0	0	83.3	16.7
Thompson matched	8	100.0	62.5	62.5	87.5	16.0	0	50	50
J. F. Deering matched	18	72.2	88.9	44.4	88.9	12.0	16.7	33.3	50
Total (weighted) Matched	106	77.0	84.9	60.4	89.6	14.4	9.4	48.7	41.8

ISFP YOUTH	**Number Complete (N)**	**Female (%)**	**Student Grade**		
			6th	**7th**	**8th**
Broad Rock matched	12	58.3	100.0	0.0	0.0
Calcutt matched	9	50.0	44.4	22.2	33.3
Park View matched	35	37.1	77.1	8.6	14.3
Curtis Corner matched	8	50.0	100.0	0.0	0.0
Pier School matched	8	16.7	87.5	0.0	12.5
Portsmouth matched	5	40.0	100.0	0.0	0.0
Thompson matched	7	57.1	71.4	0.0	28.6
J. F. Deering matched	14	46.2	38.5	61.5	0.0
Total (weighted) Matched	98	43.1	74.9	13.9	11.2
Project Total (weighted)	**204**	**60.7**	**80.1**	**38.0**	**52.0**

Fidelity

We tracked the fidelity of implementation of the three programs with checklists completed by the prevention educator after each session of the curricula. These checklists were designed to reflect the activities specified by the curriculum for each session. See Appendix D for an example. Table 4 summarizes information on the educator-reported fidelity of implementation of the curricula using the post-session process checklists we provided. The first column (Year 1 sessions) of the ISFP curriculum contains the total percent of completion of the content calculated from all data for the initial 7 sessions over the three years of the project. These percentages combine data collected across all of the sessions for parents, children, and families. The ISFP curriculum was very high in fidelity across all of the sites (95%), suggesting that it worked well in terms of feasibility of implementation. The families who attended received a uniformly high "dose" of the program. For the classroom curricula, we aggregated across the entire set of sessions for each classroom, and then averaged across all the classrooms in each school. The classroom curricula were implemented with more variability across school, and it also appears that PN was likely to be more fully implemented on average (91.5%) than LST (77.3%) in year 1. This is consistent with the higher number of required sessions for LST (15 in the first year, vs. 7 for PN), and has been reflected in conversations during the monthly liaison meetings. In year 2 the PN curricula increased to eight lessons and the percentages dropped substantially (84.1%). In year 2 the LST curricula decreased to ten lessons, but there was difficulty with scheduling classes and overlap with other programs

Table 4. Program Fidelity (Percent of Content Covered)

Program Fidelity %

	ISFP		Project Northland			Life Skills Training		
School	Year 1 sessions	Booster sessions	Year 1	Year 2	Year 3			
Broad Rock	95.0*	90.1*	95.6	79.7	95.3			
Park View	90.8		79.7	83.6	83.7			
Curtis Corner	95.0*	90.1*	95.9	81.9	92.4			
J F Deering	96.3	93.7	94.7	91.3	86.9			
Total	**93.6**	**93.7**	**91.5**	**84.1**	**89.6**			
Grand Total	**93.6**			**88.4**				
	Year 1	Booster				Year 1	Year2	Year 3
Calcutt	86.2					80.9	75.1	84.8
Pier School	94.0	99.1				83.1	73.3	76.1
Portsmouth	96.0					75.3	64.8	98.2
Thompson	97.5					69.9	80.8	99.3
Total	**93.4**	**99.1**				**77.3**	**73.5**	**89.6**
Grand Total	**96.3**						**80.1**	
	Family-based program			**School-based program**				
Project Total	**94.9**			**84.3**				

*** Fidelity compiled across the two South Kingstown schools**

and adherence to the program was compromised (73.5%). In year 3, fidelity in PN (89.6%) and LST (89.6%) schools improved resulting in the average participant in both programs receiving nearly 85% of the program across the three years.

Description of the monthly liaison meetings

The contracted role of the CRST was to coordinate a quantitative evaluation of the effectiveness of the three programs selected for implementation by the vendor agencies. In our role as the statewide evaluation team for the project, the CRST held monthly meetings over the duration of the project with staff liaisons from each of the five vendor agencies. Although the ostensible purpose of our evaluation was to draw statistical conclusions regarding program effectiveness, we found that another valuable type of learning took place at these liaison meetings. At these meetings, staff from the five participating agencies developed a collaborative style, working together to solve mutual and individual agency problems as they arose over the duration of the project. Through that process, the evaluation team has come to know a great deal more at an informal, qualitative level about the challenges of engaging in science-based prevention. The attending agency staff agree that this was an extremely useful process, lending somewhat to an action research approach to science-based prevention. The minutes from these meetings are provided in Appendix E.

Initially these meetings were held to identify shared objectives and provide trainings in program evaluation. The evaluation liaisons' functions were described as being responsible for identifying and fulfilling school policies regarding parental and child consent, responsibility for distribution and collection of evaluation measures to and from program educators, monitoring the experience of program educators with the evaluation and consulting with the statewide evaluation team on any difficulties that might arise, and responsibility for communicating summary information produced by the CRST to agencies and program educators. Typical meeting agendas discussed confidential vs. anonymous surveys in context of choosing between active and passive consent, development of a list of codes for our matching procedure, discussions of IRB policies, updates regarding program trainings and coordination of evaluation trainings, and discussions of sharing intermediate and outcome objectives for all three of the prevention curricula.

These initial meetings and a pilot of the programs assisted in refining the measures and determining the best steps for the *Replicating Science-based Substance Abuse Programs* demonstration project evaluation. Once the programs were in place, there was a growing confidence in the ability of agencies to assist in the evaluation and the focus of the agenda topics evolved into discussions of implementation issues. The CRST's role shifted from one of training and educating to one that provided technical assistance and support in implementation issues. However, the role of the liaison also switched to becoming the experts for many of the issues, as several of the liaisons were experienced program delivery staff and others were experienced supervisory staff. This mixture of supervisory, delivery, and evaluation specialists proved very effective for overcoming a variety of obstacles to implementation.

Some of the common themes in these liaison meetings involved issues with the Problem Based Prevention System (PBPS), recruiting families into the *Iowa Strengthening Families Program*, handling troublesome youth in class, establishing the best practices for getting youth to assent to taking the surveys, scheduling, fidelity, and difficulties from school and agency staff turnover. Three products resulted from these meetings. First, the agencies and the CRST worked together to create two documents that indicate how LST and PN curricula fit with the Standards of Health Curriculum in Rhode Island. These tools are invaluable for promoting the selected science-based programs into the school system (see Appendix F). A third product resulted from a focus group conducted in the spring of 2004, culminating in a report titled *Understanding the Process of Science-based Prevention: Implementer Perspectives: Report of Themes*

and Recommendations (see Appendix G). The findings from this meeting informed of common barriers and promising practices for building relationships with schools, training in science-based curricula, recruiting participants into family programs, and implementation of science-based programs.

Training in Curriculum Delivery

Agency staff who had been assigned to deliver the three curricula were trained by designated representatives of the developers during the spring of 2002. We conducted brief post-training surveys for each of the three trainings (Surveys are included in Appendix H). Table 5 provides information on a few of the items in that survey. For the most part respondents to our survey were those who were paid staff on this project, but there were others who also took the Iowa Strengthening Families training, primarily additional staff from the same agencies. Sample sizes are small and we will simply report some descriptive findings. The participants reported a moderate amount of previous experience with science-based curricula (40% to 60% had had some). The trainings were all rated relatively positively, although qualitative comments were more mixed for the Project Northland training. Ratings of confidence in one's ability to deliver the curriculum as designed, and in the curriculum's effectiveness for the particular target populations of local agencies, were a bit lower for the Life Skills Training curriculum.

Table 5. Educator Training Evaluation Ratings

Training Evaluation Item	Program Training Ratings		
	ISFP	Project Northland	Life Skills Training
Number of respondents	21.0	5.0	7.0
White (%)	71.4	100.0	71.4
Previous experience with science-based curricula (% yes)	40.0	60.0	42.9
Rating of the training: "How well did it work for you?" on a 5-point scale	4.2* "very well"	4.0** "very well"	3.8*** "very well"
Confidence in ability to deliver the program on a 5-point scale	4.3 "confident"	4.8 "very confident"	3.8 "confident"
Confidence in the curriculum's effectiveness for "your students/famiilies"	3.8 "confident"	3.8 "confident"	3.3 "moderately confident"

* Positive comments: "excellent presenters, thorough, walked us through all aspects of the program"

** Mixed comments: "needed more focus on logistics, timeline, specific tasks," "the training was clear and well presented"

*** Positive comments: "much information," "the trainer was very enthusiastic and helpful"

Sustainability

Likelihood of Institutionalizing Science-based Programs

We conducted an informal survey of the staff who had delivered the prevention curricula in the project, asking about their views of the curricula and their perception of the likelihood that the curricula would continue to be delivered at the sites where the project was located. All PN schools indicated that they planned to continue the PN curriculum. One of the four LST schools was confident they would do so, and another was uncertain.

Reflections on Potential for Institutionalization

LST has substantially more sessions than PN in the initial year, and the booster sessions for LST are very repetitive of the first year's content. Classroom teachers and agency staff who were delivering the curricula appeared to find PN a more positive experience, especially in the 2nd and 3rd years. At the end of the program those prevention educators rated PN as more appropriate for their local conditions, and more likely to be locally effective.

A second aspect of sustainability concerns staffing. Using classroom teachers employed by the school (i.e. health education teachers) to deliver the curriculum (with appropriate training from the program developers) was associated with a higher rate of institutionalization.

ISFP was viewed as an excellent program, but is very expensive and "labor intensive."

Program Outcomes

Instrumentation

Six different instruments were assembled for measuring outcomes for the evaluation of this project and can be viewed in Appendix B along with their accompanying codebooks.

Youth Participant Questionnaires.

Two surveys containing six sections were created and used for pretest and posttest measures. The first five sections made up a "standard item set" developed to function as a core for our youth self-report outcome measures across programs. This item set included tracking items, demographic and substance use items, and major risk and protective scales. Section six was devoted to measure program specific outcomes. The questionnaires were designed with the use of Teleforms© software to facilitate data entry. Program educators administered these instruments at the first session of the classroom curricula (after participant assent was obtained) and again following the last session. The same administration process was followed in years two and three. Two additional surveys identical to the surveys described above (excluding the tracking questions) were utilized to measure responses from the 8th graders in the same schools in year one of the program. These measures were utilized to provide comparison group data and were only collected on one occasion. Codebooks were created to provide the source of the items or scales used, a brief description of the item or scale, values of the response set, the procedure for creating the scale, and the objectives sought.

Table 6 contains information on the scales utilized in the LST version, which includes the source, direction of scoring and scale psychometrics. Items in section 2, and scales in sections 3 through 5 were selected for their value in providing statewide planning data and for comparability with other accessible data. Measures were adapted from various sources such as the Monitoring the Future survey (MTF; Johnston, O'Malley & Bachman, 2000), Youth Tobacco Survey (The Office of Health Statistics, Rhode Island Department of Health), Communities That Care (CTC; Hawkins & Catalano), and the Life Skills Health Survey (LSHS; Institute for Prevention Research, Department of Public Health, Cornell University Medical College; Kenneth Griffin, Gilbert Botvin). Some of the scales were renamed in order to be consistent with the scoring direction, likewise, other scales were recoded in order for higher values to indicate higher levels of a certain skill or cognitive schema (e.g., attitude). Section six was devoted to measuring program specific objectives for the Life Skills Training curricula. Drug refusal skills were measured at two conceptual levels, saying no to offers of each drug and a variety of options (e.g., changing the subject) for saying no. These two scales were combined to measure an overall drug refusal skill. Because of the low instances of substance use and the associated risk factors for the majority of youth in the program at baseline, some scales required transformations. One transformation technique used was dichotomizing, which involves changing the distribution of data. For example: changing a 7-item response (i.e., 1 = never used drugs to 7 = 40 or more occasions) to a 2-item response (i.e., never used drugs vs. one to several occasions of use). Other transformations included using the natural logarithm and adding 1 to the means of a scale in place of the original mean or, to increase the means exponentially.

Table 7 contains the scale psychometrics for year 1 of the PN survey. The PN version was identical to the LST survey in the first 5 sections, however reflects the first year administration in the schools implementing PN. Section 6 contains scales specific to the Project Northland curricula. Scales were adapted from The Partnerships for Youth Health Student Survey (PN; Cheryl Perry, Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis; Williams et al., 1995) and the CTC survey (Hawkins & Catalano). Parent communication was measured with 4 items, however one

item was excluded to improve the alpha level. The Reasons Not to Use Alcohol items have been used in other research and provided reasonable alpha. The tables of scale psychometrics for the 8th grade survey is located in Appendix I.

Iowa Strengthening Families Questionnaires.

Two surveys were created (using Teleforms© software) for the participants in the ISFP. Participants completed the same survey for both pretest at the initial session and posttest at the end of the final session. The instruments and codebooks can be viewed in Appendix B. Parents and youth were measured on the same constructs; however, the number of items was reduced for the youth version. Similar to the Youth Participant Questionnaires described above, respondents completed a unique identifier code that was used to match participant responses from pretest to posttest.

Table 8 contains the scale properties for the initial 7 sessions of the ISFP. Measures were adapted from various sources such as: Project Family researchers (Spoth, Redmond, Haggerty and Ward, 1995), Iowa Youth and Families Project (Conger, 1989), Communities That Care (Hawkins & Catalano), and the Youth Tobacco Survey (The Office of Health Statistics, Rhode Island Department of Health). The scale alphas ranged from .669 to .938, which indicates a reasonable reliability for the measure overall. One scale, the Parental Expectations of Non-use required an exponential transformation.

Table 6 Measures and Scale Psychometrics for Year 1 - LST program

Section and Scale	Source	Range	Scoring direction	Number of items	N	MIN	MAX	MEAN	STD	Alpha	Transformation type
2: Age of initiation alcohol	MTF	1= never 2= 14 years or older... 8= 8 years old or younger		1	668	1	7	1.3	0.996	N/A	dichotomized 0= never 1= initiated
2: Age of initiation tobacco	YTS	1= never 2= 14 years or older... 8= 8 years old or younger		1	670	1	7	1.1	0.726	N/A	dichotomized 0= never 1= initiated
2: Age of initiation marijuana	YTS	1= never 2= 14 years or older... 8= 8 years old or younger		1	687	1	7	1.1	0.648	N/A	dichotomized 0= never 1= initiated
2: 30 day alcohol prevalence	MTF	1= 0 occasions 7= 40 or more occasions		1	676	1	6	1.1	0.369	N/A	dichotomized 0= never 1= one or more occasion
2: 30 day tobacco prevalance	YTS	1= 0 days 7= All 30 days		1	678	1	7	1.0	0.343	N/A	dichotomized 0= never 1= one or more occasion
2: 30 day marijuana prevalance	YTS	1= 0 occasions 7= 40 or more occasions		1	675	1	7	1.0	0.292	N/A	dichotomized 0= never 1= one or more occasion
2: Problem drinking	MTF	1= none 6= 10 or more times		2	636	0	1	0.0	0.142	N/A	dichotomized 0= never 1= one or more occasion
3: Peer disapproval scale	MTF	1= Approve 3= Disapprove	Higher scores = more peer disapproval	4	688	1	3	2.9	0.389	0.94	Exponential
3: Perceived risk scale	MTF	1= No risk 3= Great risk	Higher scores = greater perceived risk	5	638	1	3	2.7	0.448	0.86	Exponential
3: Favorable attitudes toward drug use	CTC	1= Very wrong 4= Not wrong at all	Higher scores = more favorable attitudes toward substance use	4	686	1	4	1.2	0.478	0.88	Natural log plus 1
4: Family attachment scale	CTC	1= NO! 4=YES!	Higher scores = stronger family attachment	4	675	1	4	3.2	0.741	0.81	
4: Parental attitudes toward drug use	CTC	1= Not wrong at all 4= Very wrong	Higher scores = greater parental disapproval toward substance use	3	680	1	4	1.1	0.318	0.82	Natural log plus 1
5: Drug use intentions	LSHS	1= Definitely not 5= Definitely will	Higher scores = greater intentions to use within the next year	5	681	1	5	1.2	0.502	0.89	Natural log plus 1
5: Peer normative beliefs	LSHS	1= None 5= All or almost all	Higher scores = perception that more peers use drugs	5	681	1	5	2.0	0.872	0.93	
6: Drug refusal skills alternate scale	LSHS	1= Definitely would not 5= Definitely would	Higher scores = more or better refusal skills	10	661	1	5	3.9	1.194	0.91	
6: Drug attitudes - pro drug composite scale	LSHS	1= Strongly disagree 5= Strongly agree	Higher scores = positive attitude toward substance use	8	668	1	5	1.3	0.632	0.95	Natural log plus 1
6: Social skills scale	LSHS	1= Strongly disagree 4= Strongly agree	Higher scores = better social skills	5	675	0	10	2.5	2.123	0.78	
6: Stress management skills	CSAP	1= Strongly dsagree 4= Strongly agree	Higher scores = better stress management skills	4	660	0	10	3.2	2.255	0.85	

Source: MTF Monitoring the Future; YTS Youth Tobacco Survey; CTC Communities That Care; LSHS Life Skills Health Survey

Table 7 Measures and Scale Psychometrics for Year 1 - PN program

Section and Scale	Source	Range	Scoring Direction	Number of items	N	Min	Max	Mean	STD	Alpha	Transformation type
2: Age of initiation alcohol	MTF	1= never 2= 14 years or older... 8= 8 years old or younger		1	717	1	7	1.3	0.958	N/A	dichotomized 0= never 1= initiated
2: Age of initiation tobacco	YTS	1= never 2= 14 years or older... 8= 8 years old or younger		1	720	1	7	1.1	0.561	N/A	dichotomized 0= never 1= initiated
2: Age of initiation marijuana	YTS	1= never 2= 14 years or older... 8= 8 years old or younger		1	719	1	7	1.0	0.423	N/A	dichotomized 0= never 1= initiated
2: 30 day alcohol prevalence	MTF	1= 0 occasions 7= 40 or more occasions		1	714	1	7	1.0	0.301	N/A	dichotomized 0= never 1= one or more occasion
2: 30 day tobacco prevalence	YTS	1= 0 days 7= All 30 days		1	720	1	4	1.0	0.158	N/A	dichotomized 0= never 1= one or more occasion
2: 30 day marijuana prevalence	YTS	1= 0 occasions 7= 40 or more occasions		1	714	1	2	1	0.037	N/A	dichotomized 0= never 1= one or more occasion
2: Problem drinking	MTF	1= none 6= 10 or more times		2	698	0	1	0.0	0.084	N/A	dichotomized 0= never 1= one or more occasion
3: Peer disapproval scale	MTF	1 = Approve 3 = Disapprove	Higher scores = more peer disapproval	4	725	1	3	2.9	0.382	0.94	Exponential
3: Perceived risk scale	MTF	1 = No risk 3 = Great risk	Higher scores = greater perceived risk	5	680	1	3	2.7	0.375	0.79	Exponential
3: Favorable attitudes toward drug use	CTC	1 = Very wrong 4 = Not wrong at all	Higher scores = more favorable attitudes toward substance use	4	718	1	4	1.2	0.396	0.86	Natural log plus 1
4: Family attachment scale	CTC	1 = NO! 4 =YES!	Higher scores = stronger family attachment	4	713	1	4	3.3	0.656	0.79	
4: Parental attitudes toward drug use	CTC	1 = Not wrong at all 4 = Very wrong	Higher scores = greater parental disapproval toward substance use	3	721	1	4	1.1	0.299	0.89	Natural log plus 1
5: Drug use intentions	LSHS	1 = Definitely not 5 = Definitely will	Higher scores = greater intentions to use within the next year	5	718	1	5	1.1	0.353	0.80	Natural log plus 1
5: Peer normative beliefs	LSHS	1 = None 5 = All or almost all	Higher scores = perception that more peers use drugs	5	719	1	5	2.0	0.803	0.92	
6: Parent communication	PN	1 = False 2 = True	Higher scores = better parent-child communication	3	720	1	2	1.9	0.285	0.67	Exponential
6: Reasons Not to Use Alcohol	PN	1 = Not too important for me 5 = Very important for me	Higher scores = more importance to reasons and consequences for not using alcohol	10	697	1	5	4.3	0.838	0.89	Exponential
6: Interaction with antisocial peers	PN	1 = None of my friends 5 = 4 of my friends	Higher scores = more association with negative peers	6	696	1	5	1.1	0.334	0.76	Natural log plus 1
6: Rules and consequences against drinking	PN	1 = False 2 = True	Higher scores = more family rules	1	716	1	3	1.4	0.775	N/A	

Source: MTF Monitoring the Future; YTS Youth Tobacco Survey; CTC Communities That Care; LSHS Life Skills Health Survey; Partnerships for Youth Health Student Survey

Table 8 Measures and Scale Psychometrics for First 7 Sessions ISFP

	Section and Scale	Source	Range	Scoring direction	Number of items	N	MIN	MAX	MEAN	STD	Alpha	Transformation type
Parent/Caregiver	Parental limit setting and monitoring	Spoth	1 = Always 7 = Never	Higher scores correspond to positive parenting skills	13	145	3.4	6.5	5.4	0.623	0.73	
	Family nurtruing and protective qualities	Spoth	1 = Always 7 = Never	Higher scores correspond to positive parenting skills	7	145	2.8	6.9	5.3	0.852	0.80	
	Parent-child communication	Spoth	1 = Strongly disagree 5 = Strongly agree	Higher scores correspond to positive parenting skills	5	142	2.2	5	4.2	0.618	0.82	
	Parent-child involvement	Spoth	1 = Not true 4 = Always or almost always	Higher scores correspond to positive parenting skills	6	142	2.2	4	3.2	0.457	0.74	
	Parental expectations of non-use	Spoth	1 = Strongly disagree 5 = Strongly agree	Higher scores correspond to positive parenting skills	3	140	1	5	4.4	0.933	0.87	Exponential
Youth	Parental limit setting and monitoring	CTC	1 = NO! 4 = YES!	Higher scores correspond to higher preceived parenting skills	6	96	2.3	4	3.5	0.408	0.67	
	Family nurtruing and protective qualities	Spoth	1 = Always 7 = Never	Higher scores correspond to higher preceived parenting skills	7	98	2	7	5.5	1.175	0.81	
	Parent-child communication	FFS	1 = Never 7 = Always	Higher scores correspond to higher preceived parenting skills	8	94	1	7	4.2	1.441	0.87	
	Parent-child involvement	Spoth	1 Not true 4 = Always or almost always true	Higher scores correspond to higher preceived parenting skills	6	94	1	4	3.1	0.685	0.89	
	Parental expectations of non-use	YTS	1 = Strongly disagree 4 = Strongly agree	Higher scores correspond to higher preceived parenting skills	3	92	4	4	3.7	0.673	0.94	Exponential

Source: Spoth Project Family; Iowa Youth and Families; CTC Communities That Care; FFS Family Functioning Scale

Major Questions Addressed: Outcomes on Intermediate Objectives in Year 1

Intermediate PN Objectives for Year 1

Survey data from the pre-test and post-test in Year 1 were used to examine the four intermediate objectives established for Project Northland. As Table 9 indicates, the combined results for the four sites yielded extremely small changes and none of the individual sites receiving PN achieved any of these objectives.

Intermediate LST Objectives for Year 1

Table 9 also reports the findings for the LST intermediate objectives in Year 1. One of the five intermediate objectives was achieved when data from the fours sites were combined: There was a 19.5% increase in self-rated social skills from the pre-test to the post-test. Three of the four sites were very successful on this objective. None of the other LST objectives was achieved in Year 1 for the combined sites, although one site did achieve a 24% increase in stress management skills.

Intermediate ISFP Objectives for Year 1

The sample sizes for individual sites were too low to make comparisons for each site, so only the combined data for all eight sites are reported in Table 9. When we examined our data for both parents/caregivers and youth, none of the five objectives was achieved. The most success was for Family Nurturing, which achieved a 6.2% increase in parent ratings (the objective called for 10%). One anomalous finding was the contrast between parents' report of a 6% increase in Parental Expectations for Non-use, while youth reported a 12% decrease.

Major Questions Addressed: Outcomes on Intermediate Objectives Over Three Years

Eighth Grade Treatment Group vs. Eighth Grade Comparison Group on Intermediate Objectives

Shared Intermediate Objectives

Eighth graders who participated for the three years of the two classroom programs were compared to "untreated" eighth graders who were surveyed in the spring of the year the programs began (2002). For the seven scales measuring intermediate outcomes that were administered to both the LST and the PN sites, we examined the results for all of the participants with a MANOVA. The results are summarized in Table 10. The F-test for the MANOVA was significant, indicating that the combination of the scales significantly differentiated between the treatment group (eighth graders who had been in the program for three years) and the comparison group (eighth graders who had completed the survey in the year the program began, and did not receive the program). Follow-up analyses indicated that in all cases the scale scores showed better results for the treated group, and in all but one case these differences were highly significant ($p<.001$). The largest differences were for Favorable Attitudes toward Drug Use (the treatment group had 14.5% lower favorable attitudes) and Drug Use Intentions (the treatment group had 16.4% lower intentions to use drugs). The only non-significant difference was for the Family Attachment Scale, a variable only peripherally related to the two in-school curricula.

Table 9 PERCENT CHANGES ON INTERMEDIATE OBJECTIVES IN YEAR ONE

Objectives	Positive=	Site 1	Site 2	Site 3	Site 4	Weighted Combination	Met Objective
Project Northland							
Favorable attitude toward alcohol use	10% decrease	1.20%	-2.80%	0.70%	-0.20%	**-0.23%**	
Involvement with antisocial peers	10% decrease	8.00%	1.90%	1.00%	-4.20%	**1.19%**	
Parent communication re alcohol	25% increase	-0.50%	1.60%	1.60%	0.50%	**0.76%**	
Rules and consequences re drinking	10% increase	-1.10%	0.00%	0.50%	0.50%	**-0.02%**	
Life Skills Training							
Social Skills	20% increase	26.00%	3.40%	28.10%	59.60%	**19.50%**	**yes**
Drug Refusal Skills	20% increase	9.90%	1.00%	5.40%	-6.40%	**5.44%**	
Stress Management Skills	20% increase	-1.90%	-17.40%	-8.10%	24.20%	**-7.93%**	
Favorable Attitudes toward Drugs	20% decrease	-4.00%	2.30%	-4.70%	-8.60%	**-2.09%**	
Perceived Peer Use	25% decrease	7.50%	2.10%	7.10%	1.50%	**5.40%**	
Iowa Strengthening Families Program							
Parent/Caregiver							
Parental Limit Setting	20% increase					**1.60%**	
Family Nurturing & Protective	10% increase					**6.20%**	**~**
Parent/child Communication	30% increase					**1.90%**	
Positive Parent/child Involvement	20% increase					**3.30%**	
Parental Expectations for Non-use	10% increase					**5.70%**	
Youth							
Parental Limit Setting	20% increase					**3.40%**	
Family Nurturing and Protective	10% increase					**2.50%**	
Parent/child Communication	30% increase					**7.70%**	
Positive Parent-child Involvement	20% increase					**3.30%**	
Parental Expectations of Non-use	10% increase					**-12.10%**	

Table 10. Comparisons of Eighth Grade Students (Comparison Group vs. Treatment Group) Shared Scales						
Scale	Range	Best Possible Score	Comparison Group	Treatment Group	Difference	*p*-value
Peer disapproval	1-3	3	2.7	2.8	3%	.001
Perceived risk	1-3	3	2.6	2.7	2%	.001
Favorable attitudes toward drug use	1-4	1	1.8	1.5	-15%	.001
Family attachment	1-4	4	2.9	2.9	0.3%	.424
Parental attitudes toward drug use	1-4	4	3.7	3.8	3%	.001
Drug use intentions	1-5	1	1.7	1.4	-16%	.001
Peer normative beliefs	1-5	1	2.5	2.3	-8%	.001

Non-shared Intermediate Objectives

For measures of intermediate objectives that were non-shared (that is, measures based on items that were only completed by students receiving one or the other program), we examined the amount and significance of the difference between the two groups. Table 11 provides these results, which are summarized below.

Project Northland only

The F-test for the MANOVA summary index is significant, indicating that there is evidence for group differences on a linear combination of the dependent variables. In other words, when all of the intermediate non-shared intermediate outcomes are looked at together there is a statistically significant difference between the treatment group and the comparison group. The multivariate effect size, or eta-squared (1- Wilks' Lambda) is small to moderate (.06). The F-tests for the follow-up ANOVAs on the dependent variables are all significant ($p<.01$) as well, with mostly small effect sizes. Mean comparisons indicate that the treatment group significantly outperformed the comparison group for each non-shared intermediate objective. The magnitudes of the differences are relatively small, ranging from an 8% higher negative attitude toward alcohol use for the treatment group, to 7% higher perceived quality of parent communication, to 6% lower reported interaction with negative peers and 6% higher perception that parents enforced rules and had consequences for ATOD use.

Life Skills Training only

The F-test for the MANOVA summary index is significant. This indicates that the treatment and the comparison group differ on a linear combination of the dependent variables. As for the PN comparison, when all of the non-shared intermediate outcomes are looked at together, the treatment group did significantly better than the comparison group. The multivariate effect size, or eta-squared (1- Wilks' Lambda) is small to moderate (.09). Of the five F-tests for the follow-up ANOVAs, three are significant. The treatment group had 8% higher drug refusal skills ($p=.000$), 18% lower pro-drug attitudes ($p<.000$), and 6% lower perceived peer norms for ATOD use ($p=.000$). Effect sizes are in the small to moderate range. The differences for social skills and stress management were not significant.

Table 11. Comparisons of Eighth Grade Students (Comparison Group vs. Treatment Group) Non-shared Scales					
Items	Range	Comparison Group	Treatment Group	% Difference	*p*-value
Project Northland Only					
Negative attitude toward alcohol use	1-5	3.5	3.8	8%	.001
Interaction with antisocial peers	1-5	1.4	1.3	-6%	.003
Parent communication	1-2	1.7	1.8	7%	.001
Rules and consequences	1-2	1.8	1.9	6%	.001
Life Skills Training Only					
Social Skills Scale	0-10	7.9	8.0	1%	.142
Drug Refusal Scale	1-5	3.9	4.2	8%	.001
Stress Management Scale	0-10	7.9	7.0	-2%	.419
Drug Attitudes: Pro-drug	1-5	1.7	1.4	-18%	.001
Peer Normative Beliefs	1-5	2.4	2.3	-6%	.001

Time 1 to Time 6 Comparisons (PN Treatment Group vs. LST Treatment Group) on Selected Intermediate objectives

We ran four ANCOVAs to test whether the programs had a differential effect on two intermediate outcomes: Family Attachment (FAS) and Parental Attitudes Towards Drug Use (PATDU). Family attachment and perceived parental attitudes toward drug use can conceivably improve over time. The other scales analyzed in the 8th to 8th grade comparisons above are confounded with normal developmental changes and therefore were not analyzed. The covariate in these analyses was the intermediate outcome at time 1. The ANCOVAs indicate there were no differential program effects on an overall basis. Also there were no significant interaction effects by gender or by SES. The only marginally significant result (p=.04) is the two genders differed on family attachment with females reporting somewhat lower attachment.

Graphic displays in Figures 4, 5, 6, and 7 provide more detail on how Family Attachment and Parental Disapproval of Drug Use (as reported by the youth) changed over the three years of the intervention. Figures 4 and 5 both display consistent declines over time in Family Attachment for all of the subgroups across programs, as would be expected for early adolescents. There were no differences across the programs in the patterns over time. The two noticeable aspects of the pattern beyond the general decline are somewhat lower attachment for non-White participants and females (the latter was significant in the previously reported ANCOVA results). Figure 6 suggests that initial positive program effects on Parental Disapproval were overcome by normal developmental decline during early adolescence. For males there appeared to be an initial difference between the programs, with LST males reporting lower Parent Disapproval than PN males, followed by a greater increase in disapproval for the LST males. However, this pattern was not statistically significant. Figure 7 shows a consistent decline over time for both White and non-White groups after the initial sessions of the curricula and hints at White youth perceiving more parental disapproval than their non-White peers over the three years. However, these differences were not statistically significant. There was a large increase in perceived disapproval after the initial 15 sessions of LST for non-White participants.

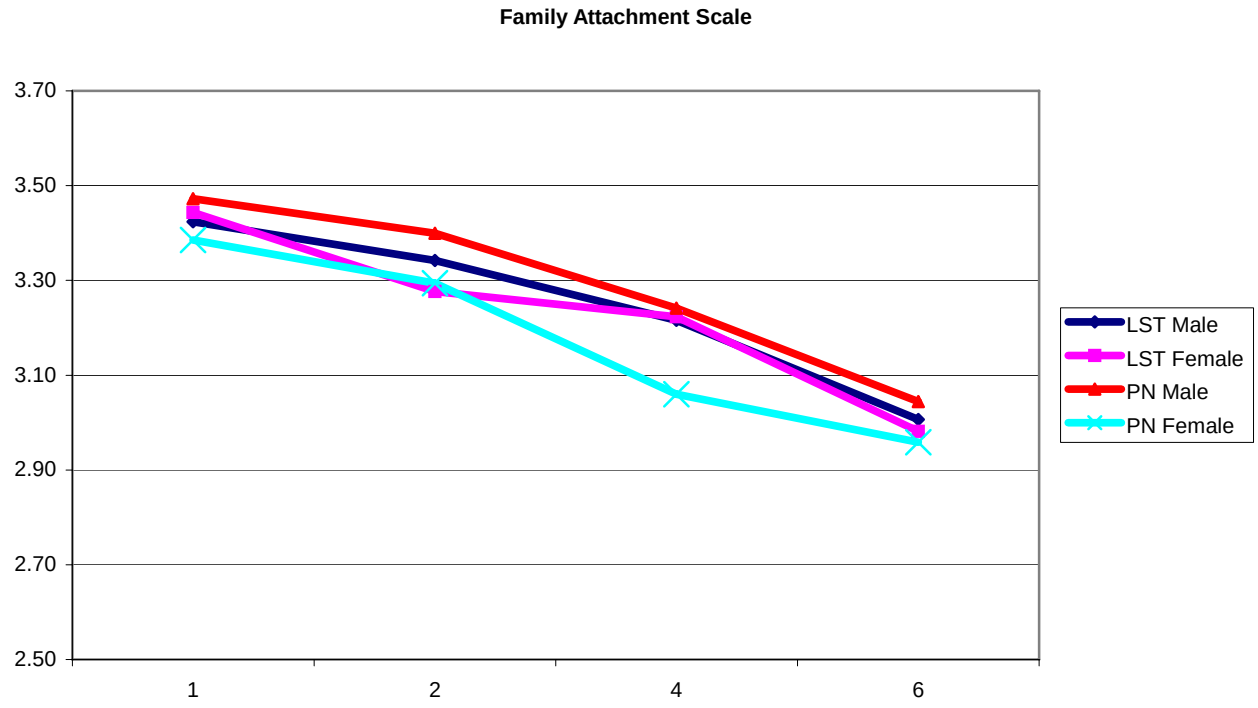

Figure 4. Family Attachment by Gender

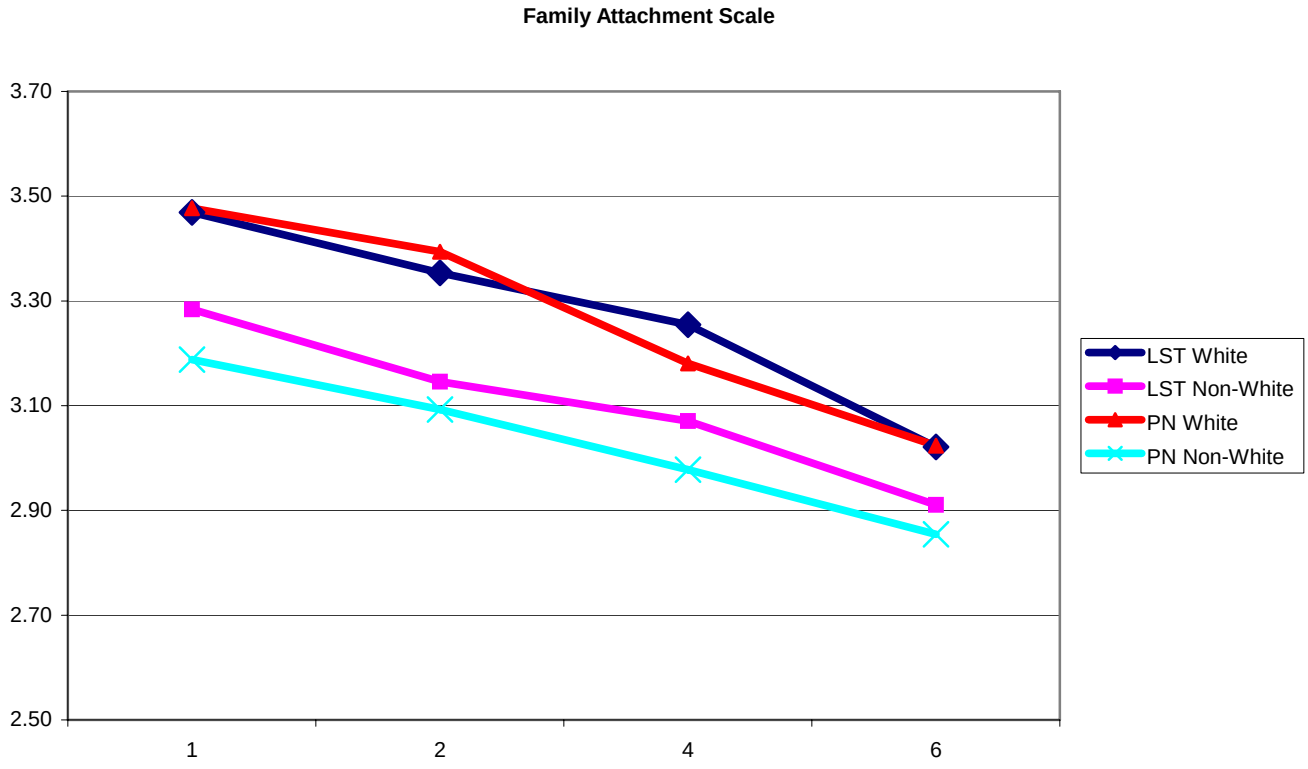

Figure 5. Family Attachment by Race

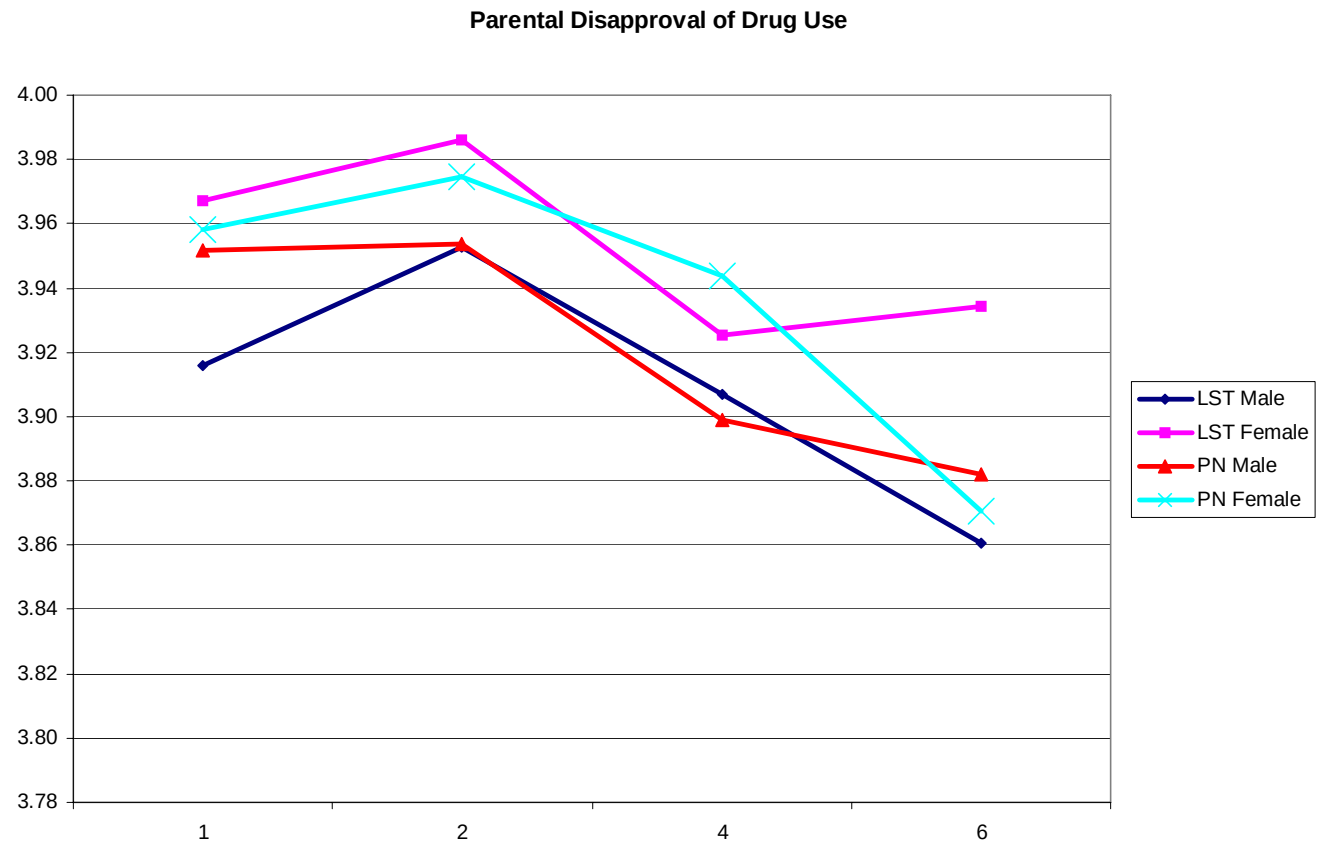

Figure 6. Parental Disapproval of Drug Use by Gender

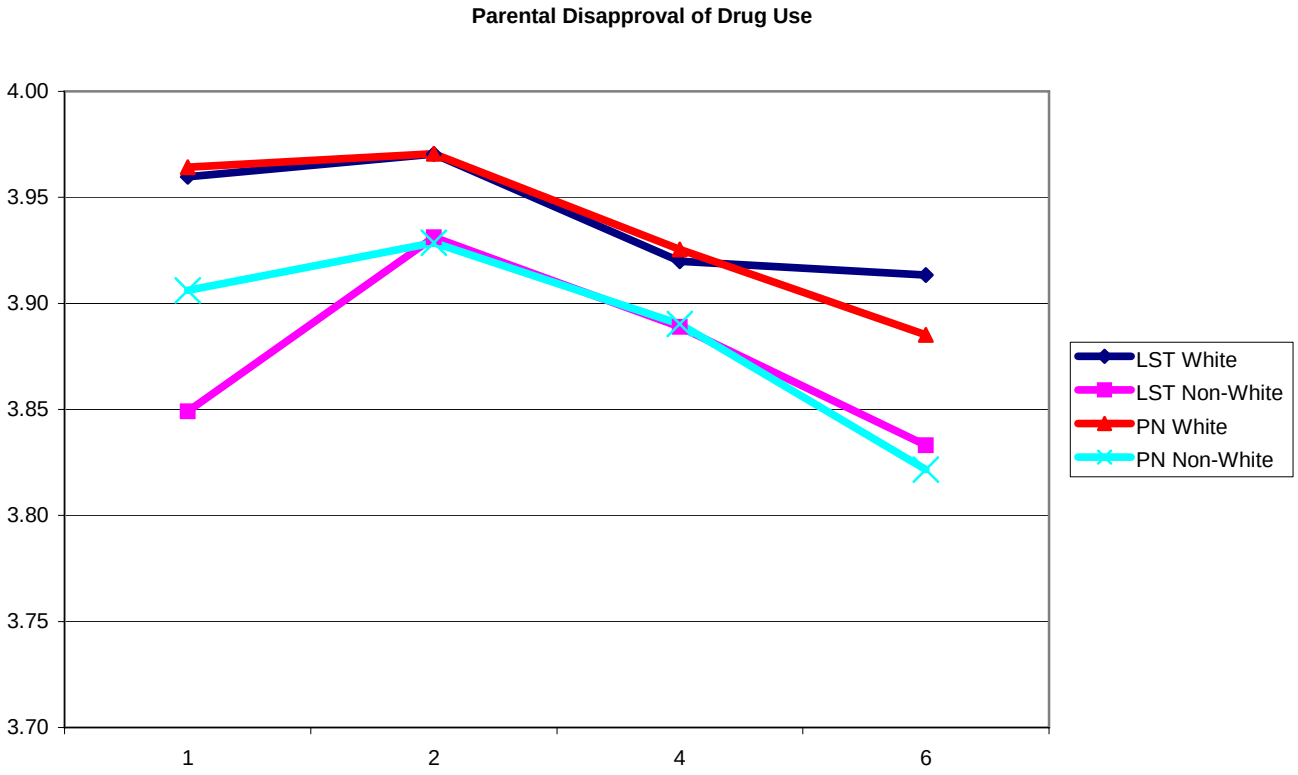

Figure 7. Parental Disapproval of Drug Use by Race

Intermediate ISFP Objectives: Statistical Comparisons

We ran paired samples t tests to determine whether there were significant changes in the intermediate outcomes for the ISFP project. For the parents, all five intermediate outcomes were significantly ($p<.01$) higher at time 2, compared to time 1 (see Tables 12 and 13). These results suggest that the program was very effective for the parents.

Table 12. Mean ratings by parents of intermediate variables in the ISFP program

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	T1 Family Nurturing & Protective Qualities	5.3	103	.823	.081
	T2 Family Nurturing & Protective Qualities	5.7	103	.632	.062
Pair 2	T1 Parent-Child Communication	4.2	102	.604	.060
	T2 Parent-Child Communication	4.3	102	.471	.047
Pair 3	T1 Parent-Child Involvement	3.2	102	.450	.045
	T2 Parent-Child Involvement	3.4	102	.412	.041
Pair 4	T1 Parental Expectations of Non-use	4.4	101	.767	.076
	T2 Parental Expectations of Non-use	4.7	101	.471	.047
Pair 5	T1 Parental Limit Setting & Monitoring	5.5	104	.618	.061
	T2 Parental Limit Setting & Monitoring	5.6	104	.542	.053

Table 13. T-tests for pre-test to post-test changes on parent-rated ISFP intermediate variables

	Mean Difference	Std. Deviation	*t*	*df*	Sig (2-tailed)
T1 Family Nurturing & Protective Qualities T2 Family Nurturing & Protective Qualities	-.360	.602	-6.06	102	.000
T1 Parent-Child Communication T2 Parent-Child Communication	-.152	.538	-2.86	101	.005
T1 Parent-Child Involvement T2 Parent-Child Involvement	-.168	.357	-4.77	101	.000
T1 Parental Expectations of Non-use T2 Parental Expectations of Non-use	-.337	.694	-4.88	100	.000
T1 Parental Limit Setting & Monitoring T2 Parental Limit Setting & Monitoring	-.160	.540	-3.02	103	.003

For the youth, however, only one difference was significant ($p=.021$) (see Tables 14 and 15). The youth at time 2 rated their parents higher than they rated them at time 1 for the Limit Setting and Monitoring variable. In the case of youth-rated parental expectations of non-use, the youth actually seemed to be indicating a decrease in expectations of non-use (although this was not significant), contrary to the parents.

Table 14. Mean ratings by youth of intermediate variables in the ISFP program					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	T1 Family Nurturing & Protective Qualities	5.5	98	1.175	.119
	T2 Family Nurturing & Protective Qualities	5.5	98	1.175	.119
Pair 2	T1 Parent-Child Communication	4.2	94	1.441	.149
	T2 Parent-Child Communication	4.4	94	1.579	.163
Pair 3	T1 Parent-Child Involvement	3.1	94	.685	.071
	T2 Parent-Child Involvement	3.2	94	.735	.076
Pair 4	T1 Parental Expectations of Non-use	3.7	92	.673	.070
	T2 Parental Expectations of Non-use	3.5	92	.959	.100
Pair 5	T1 Parental Limit Setting & Monitoring	3.5	96	.408	.042
	T2 Parental Limit Setting & Monitoring	3.6	96	.405	.041

Table 15. T-tests for pre-test to post-test changes on youth-rated ISFP intermediate variables					
	Mean Difference	Std. Deviation	*t*	*df*	Sig (2-tailed)
T1 Family Nurturing & Protective Qualities T2 Family Nurturing & Protective Qualities	.000				*
T1 Parent-Child Communication T2 Parent-Child Communication	-.171	1.219	-1.36	93	.177
T1 Parent-Child Involvement T2 Parent-Child Involvement	-.083	.538	-1.50	93	.138
T1 Parental Expectations of Non-use T2 Parental Expectations of Non-use	.201	1.083	1.78	91	.078
T1 Parental Limit Setting & Monitoring T2 Parental Limit Setting & Monitoring	-.089	.370	-2.35	95	.021

*There was no difference in scores between T1 and T2. Therefore, a significance level could not be calculated.

Major Questions Addressed: Outcomes on Ultimate Objectives Over Three Years

Eighth Grade Treatment Group vs. Eighth Grade Comparison Group

Eighth graders who participated for the three years of the two classroom programs were compared to "untreated" eighth graders who were surveyed in the spring of the year the programs began (2002). For measures of outcome objectives (i.e., 30-day prevalence of use, initiation of use during 7th and 8th grades, and problem drinking), we examined the amount and significance of the difference between the two groups. We used logistic regressions to compare the treatment and comparison groups, and it is important to note that because we calculated separate significance tests for each dependent variable we increased the risk of family-wise error. However, we believe the consistency in the pattern of findings supports their credibility. We have selected 30-day prevalence of alcohol use and initiation of alcohol use as the best illustrations of program effects. We have also selected three demographic variables, eligibility for subsidized lunch (which reflects socioeconomic status), White non-Hispanic/Latino vs. non-White and Hispanic/Latino, and gender, as control factors to examine whether the programs had differential effects on demographic sub-groups.

Use of Alcohol in the Past 30 Days

The effects of receiving a curriculum on 30-day prevalence of alcohol use were statistically significant ($p<.001$) when PN and LST schools were combined (see Table 16). Students who received a prevention curriculum were about 46% less likely to report using alcohol in the past 30 days than an untreated comparison group. There were no overall statistical differences between the effects of the two programs.

Table 16. Comparisons of Eighth Grade Students (Treatment Group vs. Comparison Group) on 30-Day Alcohol Use

Program	Treatment	Control	% Difference
LST	11%	23%	-52.2%
PN	13%	21%	-38.1%
Combined	12%	22%	-45.5%

When the effects of the two programs were looked at separately for white students and non-white students, there were differences in program effects (see Table 17). The two programs were significantly different in their effects on white students ($p<.05$), with LST having greater effects than PN on 30-day prevalence of alcohol use for white students. There were no significant differences between the programs in effects on non-whites.

Table 17. Comparisons of 8th Grade Students (Treatment Group versus Comparison Group) on 30-Day Alcohol Use, Controlling for Race

Program	Race	Treatment	Comparison	% Difference
LST	White	9%	22%	-59.1%
	Non-White	14%	28%	-50.0%
PN	White	11%	17%	-35.3%
	Non-White	18%	30%	-40.0%

When the effects of the two programs were looked at separately for students who qualified for subsidized lunches and those who did not, there were differences in program effects (see Table 18). The two programs were marginally different ($p<.06$) in their effects on full-pay students, with LST having greater effects than PN on 30-day prevalence of alcohol use for full-pay students. There were no significant differences between the programs in effects on students receiving lunch subsidies.

Table 18. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on 30-Day Alcohol Use, Controlling for S.E.S.

Program	Race	Treatment	Comparison	% Difference
LST	Full-pay Lunch	8%	21%	-61.9%
	Subsidized Lunch	15%	25%	-40.0%
PN	Full-pay Lunch	13%	20%	-35.0%
	Subsidized Lunch	14%	25%	-44.0%

When the effects of the two programs were looked at separately for male and female students, there were differences in program effects (see Table 19). The two programs were marginally different ($p<.06$) in their effects on males, with LST having greater effects than PN on 30-day prevalence of alcohol use for male students. There were no significant differences between the programs in effects on females.

Table 19. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on 30-Day Alcohol Use, Controlling for Gender				
Program	Gender	Treatment	Comparison	% Difference
LST	Male	11%	26%	-59%
	Female	12%	20%	-41%
PN	Male	13%	20%	-33%
	Female	13%	21%	-38%

Initiation of Alcohol Use During Seventh and Eighth Grade

The effects of receiving a curriculum on initiation of alcohol use were statistically significant ($p<.001$) when PN and LST schools were combined (see Table 20). Students who received a prevention curriculum were 34% less likely to report using alcohol in the past 30 days than the untreated comparison group. There was an overall statistically significant difference between the effects of the two programs, with PN substantially more likely to produce effects on initiation of use.

Table 20. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on Initiation of Alcohol Use			
Program	Treatment	Control	% Difference
LST	14%	14%	0%
PN	19%	33%	-42.4%
Combined	17%	25%	-34.0%

When the effects of the two programs on initiation of alcohol use were looked at separately for white students and non-white students, there were differences in program effects (see Table 21). The two programs were significantly different in their effects on non-white students ($p<.02$), with PN having greater effects than LST on initiation of alcohol use for non-white students. There were no significant differences between the programs in effects on white students.

Table 21. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on Initiation of Alcohol Use, Controlling for Race				
Program	Race	Treatment	Comparison	% Difference
LST	White	10%	13%	-23.1%
	Non-White	20%	23%	-13.0%
PN	White	19%	31%	-38.7%
	Non-White	17%	37%	-54.0%

When the effects of the two programs were looked at separately for students who qualified for subsidized lunches and those who did not, there were differences in program effects on initiation of alcohol use (see Table 22). The two programs were significantly different ($p<.003$) in their effects on subsidized students, with PN associated with substantial reductions in initiation of alcohol use for subsidized (lower S.E.S.) students, while LST was associated with somewhat higher initiation of use. There were no significant differences between the programs in effects on students who paid full price for lunch.

Table 22. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on Initiation of Alcohol Use, Controlling for S.E.S.				
Program	Race	Treatment	Comparison	% Difference
LST	Full-pay Lunch	10%	13%	-23.1%
	Subsidized Lunch	19%	16%	18.7%
PN	Full-pay Lunch	18%	31%	-41.9%
	Subsidized Lunch	22%	39%	-43.6%

When the effects of the two programs on initiation of alcohol use were looked at separately for male and female students, there were differences in program effects (see Table 23). The two programs were significantly different in their effects on male students ($p<.05$), with PN having greater effects than LST on initiation of alcohol use for males. There was also a significant difference in program effects on female students ($p=.001$) with PN associated with a substantial positive program effect whereas LST was associated with an anomalous finding of higher initiation by the treatment group.

Table 23. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on Initiation of Alcohol Use, Controlling for Gender				
Program	Gender	Treatment	Comparison	% Difference
LST	Male	14%	17%	-13%
	Female	15%	12%	22%
PN	Male	18%	32%	-45%
	Female	19%	33%	-43%

Problem drinking

The programs did have a significant effect ($p<.001$) on problem drinking (see Table 24). Eighth grade youth in the comparison group had a 10.9% rate of binge drinking, while the eighth grade youth in the treatment group had a 6.5% rate. This represents a 40% decrease. Analyses that compared the overall effects of the two programs and the effects of the two programs on different demographic groups did not reveal any significant ($p>.05$) differential effects.

Table 24. Comparisons of Eighth Grade Students (Treatment Group versus Comparison Group) on Problem Drinking			
Program	Treatment	Control	% Difference
LST	6.0%	12.2%	-51%
PN	7.0%	10.0%	-31%
Combined	6.5%	10.9%	-40%

Additive effects of ISFP participation on alcohol use (30-day prevalence)

In order to address the question of the value of adding the ISFP program to the in-school curricula, we used self-report from youth participants in the classroom interventions to identify those who had also participated with their family members in the ISFP program. We examined the pattern of change in alcohol use over the three years of the project for our treatment group, comparing those who had or had

not received ISFP along with either PN or LST. Logistic regression analyses indicated that the ISFP variable (0 = did not participate, 1 = did participate) was not a significant (p>.05) predictor of alcohol initiation or 30-day alcohol prevalence. This result is not surprising, however, given the small number of students who participated in ISFP (n=98) compared to the number of students who did not participate in the program. Of the 709 youth that we were able to track across the entire three years of the program, only 21 youth received both LST and ISFP, and only 51 received both PN and ISFP.

In addition we examined these patterns graphically. We included the initial pre-test (time 1) and the three post-tests (times 2, 4, and 6). Figure 8 presents these findings. The treatment cohort is separated into four groups: LST + ISFP, LST without ISFP, PN + ISFP, and PN without ISFP. The graphs indicate that the youth who received LST or PN with ISFP had moderately better results than those youth who only received the LST or PN program. However, these results were not significant (p>.05). Despite the lack of a statistically significant difference between the groups, the pattern for the LST + ISFP group deserves comment. At the point of the initial pre-test this group was by far the most likely to have used alcohol in the past month (15% reported doing so), but by the final post-test in the third year (time 6) they reported the least use (5%). The other three groups all reported modest increases in alcohol use. The PN + ISFP group initially reported the lowest amount of use and ended with relatively low use as well. It is possible that selection factors played a role: at LST schools, higher risk families may have been more likely to participate in ISFP, while in PN schools it may have been the low risk families who were more likely to participate.

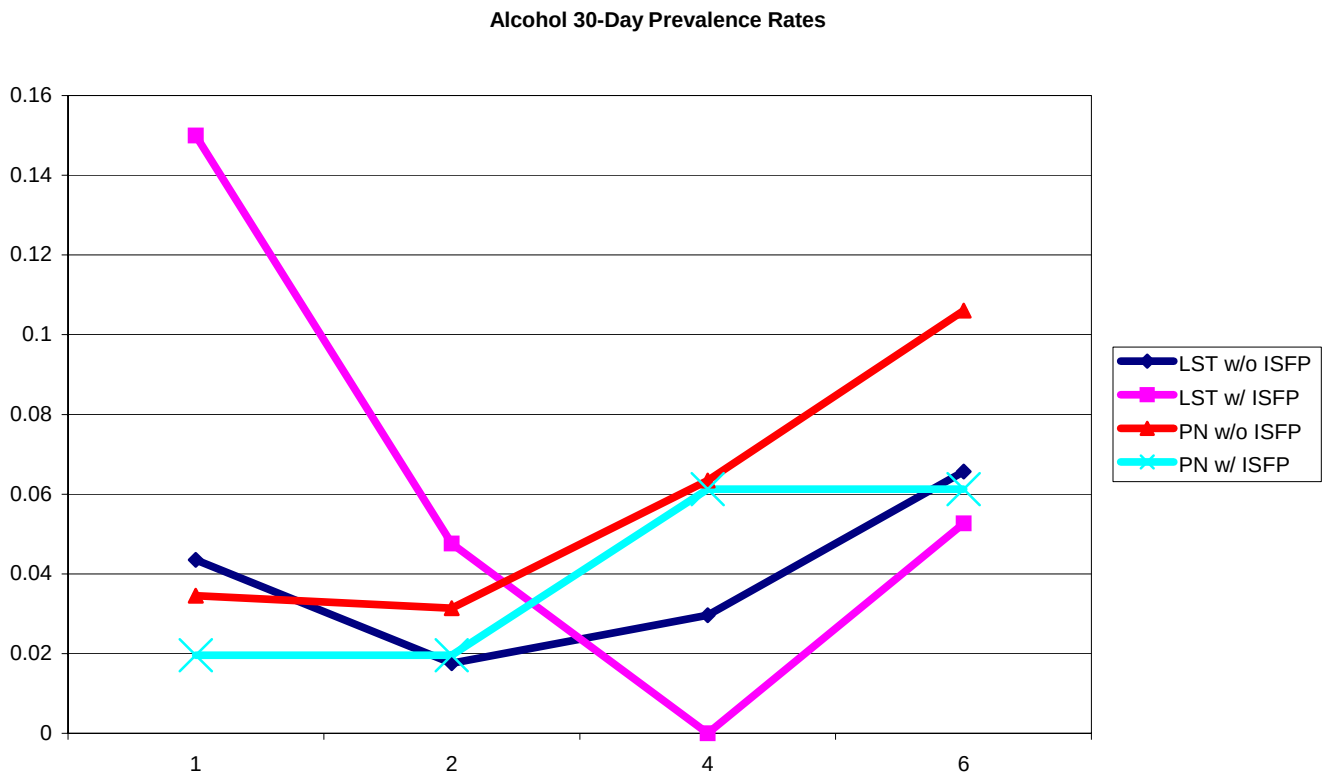

Figure 8 Comparing youth who received ISFP to those who did not

SALT Survey Results for 30-Day Prevalence of Alcohol Use

Our comparison of two groups of eighth graders, those who had received the programs and another group from the same schools who had not, showed very clear and positive effects for the programs. However, it was possible that the comparison group of eighth graders, surveyed in the spring of 2003, appeared to have a higher prevalence of alcohol use than our treatment group, surveyed in the spring of 2005, because eighth graders across the state had a higher prevalence in 2003 than in 2005. In order to check that possibility we examined SALT student survey data on 30-day prevalence of alcohol use by eighth graders for four school years: 2001-2002 through 2004-2005. SALT data, drawn from every public school in Rhode Island each year, are usually collected in the early to mid spring, similar to the timing for our survey. Figure 9 displays the results across the four years. For the first three years, the eighth graders in our treatment schools showed similar levels of alcohol use to the rest of the eighth graders in the state, starting from an almost identical level (approximately 29%), moving to slightly higher levels, and then to slightly lower levels. In the fourth year, 2004-2005, when the students who had received the programs for the past three years were in eighth grade, there is a clear downward trend not shared by the rest of the schools. The reduction in rate of use between eighth graders in 2002-2003, when our comparison group was surveyed, and 2004-2005, when our treatment group was surveyed, is 43%. This is quite similar to the 45% difference we found in our own survey (see Table 16). We conclude that the difference we found with our own survey is not a result of a statewide trend toward lower use over the course of the study, strengthening our belief that the effectiveness of the programs accounts for the difference.

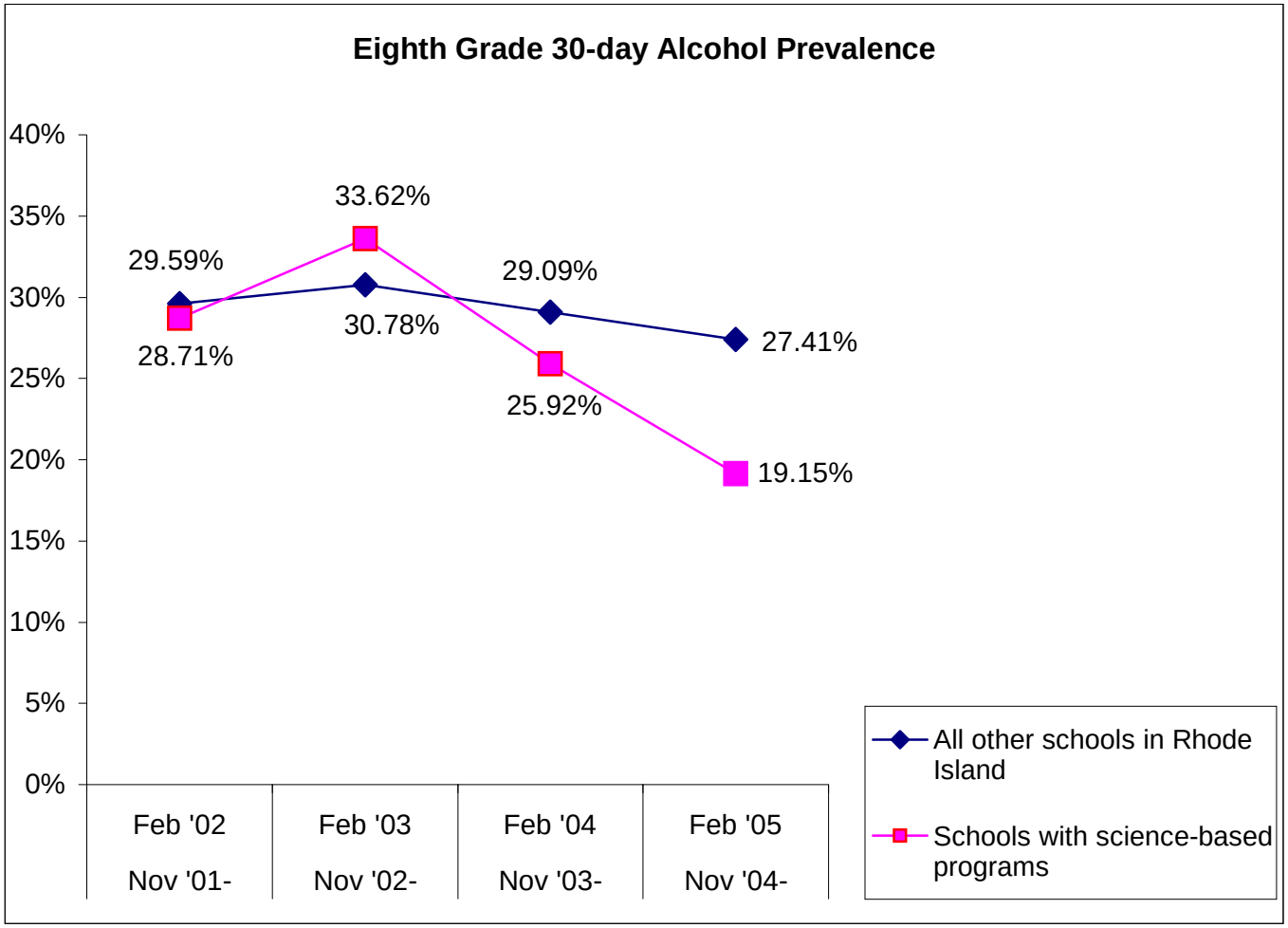

Figure 9. Comparison of Science-based treatment groups with SALT survey results over 4 years

Contrast of Sixth Grade Pre-test to Eighth Grade Post-test for the Treatment Group

For measures of outcome objectives (i.e. 30-day prevalence of use and problem drinking; initiation of use cannot be used in this type of comparison), we examined the amount of the change in use from sixth grade (initial pre-test) to eighth grade (final post-test) for the youth who we could track. We used logistic regressions to determine whether the type of program that the youth received could predict their use or non-use in eighth grade, controlling for their use levels in sixth grade. It is important to note that because we calculated separate significance tests for each dependent variable we increased the risk of family-wise error. Consistent with our report for the comparison of treatment and comparison groups, we have selected 30-day prevalence of alcohol use as the best illustration of program effects. We also selected three demographic variables, gender, eligibility for subsidized lunch (which reflects socioeconomic status), and White non-Hispanic/Latino vs. non-White and/or Hispanic/Latino, as control factors to examine whether the programs had differential effects for demographic sub-groups.

As expected, alcohol use in the past 30 days increased from sixth grade to eighth grade for youth in both the LST and PN samples (see Table 25). Although it appears that the increase in use was not as great for the LST students, we did not find that the type of program was significantly related to use after controlling for initial use.

Table 25. Comparisons of Sixth Grade to Eighth Grade Reports of 30-Day Alcohol Use by the Treatment Group

Program	6th-pre-test	8th post-test	Absolute Change
LST	5.9%	9.4%	3.5%
PN	3.4%	11.8%	8.3%
Combined	4.6%	10.7%	6.1%

Likewise, we did not find that the program had differential effects for the various demographic sub-groups (see Table 26 for race, 27 for S.E.S., and 28 for gender). However, it is interesting to note that substance use for non-whites in LST went down over the three years.

Table 26. Comparisons of Sixth Grade to Eighth Grade Reports of 30-Day Alcohol Use by the Treatment Group, Controlling for Race

Program	Race	6th Pre-test	8th Post-test	Absolute Change
LST	White	3.4%	9.6%	6.2%
	Non-White	11.6%	8.7%	-2.9%
PN	White	2.6%	11.2%	8.6%
	Non-White	7.1%	13.9%	6.8%

Table 27. Comparisons of Sixth Grade to Eighth Grade Reports of 30-Day Alcohol Use by the Treatment Group, Controlling for S.E.S.

Program	S.E.S.	6th Pre-test	8th Post-test	Absolute change
LST	Full-pay Lunch	4.5%	8.4%	3.9%
	Subsidized Lunch	8.9%	12.1%	3.2%
PN	Full-pay Lunch	2.9%	10.5%	7.6%
	Subsidized Lunch	5.8%	17.3%	11.5%

Although it appears that males receiving the LST treatment were less likely to increase their use of alcohol over time, there were no significant differential effects by gender. For one dependent variable, program did predict substance use. The type of program predicted 30-day prevalence of marijuana use, even after controlling for initial use. Consistent with the pattern of use across all substances we included in our questionnaire, the PN youth started out in sixth grade with lower levels of use but rose to higher levels of use following the three years of the program.

Table 28. Comparisons of Sixth Grade to Eighth Grade Reports of 30-Day Alcohol Use by the Treatment Group, Controlling for S.E.S.

Program	Gender	6th Pre-test	8th Post-test	Absolute change
LST	Male	7.2%	8.2%	1.0%
	Female	4.6%	9.9%	5.3%
PN	Male	3.1%	11.2%	8.1%
	Female	3.5%	12.0%	8.5%

Graphic presentation of alcohol 30-day prevalence over three years

We examined the pattern of change in alcohol use over the three years of the project for our treatment group. We included the initial pre-test (time 1) and the three post-tests (times 2, 4, and 6). Figures 10 and 11 present these findings. In Figure 10, the treatment cohort is separated into four groups: LST males, LST females, PN males, and PN females. For all four of the groups there is a predictable increase in alcohol use across the three post-tests. The only striking observation is that LST males show a large decrease in reported alcohol use between time 1 and time 2. Because of that, they ultimately have a lower use level than the other groups at the end of the third year.

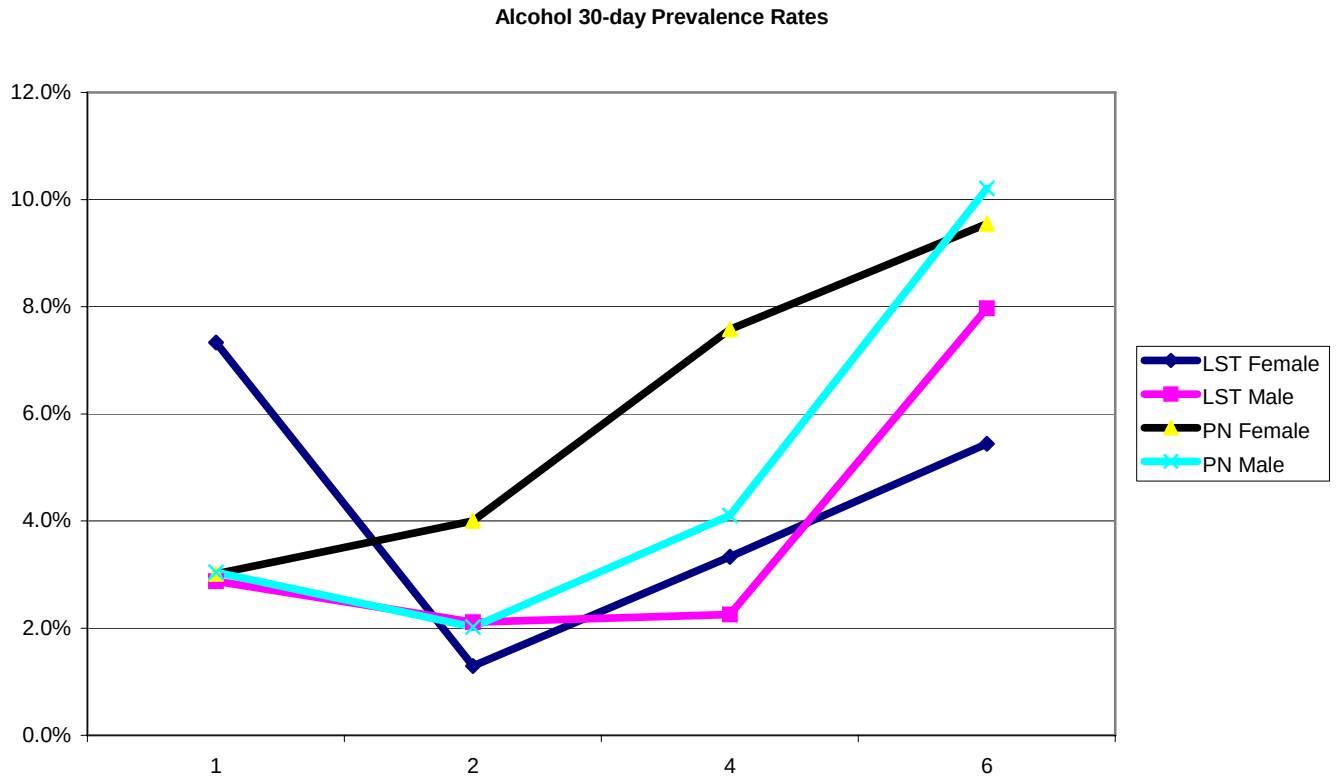

Figure 10. 30-day Alcohol Prevalence by Gender

Figure 11 separates the treatment cohort into four groups distinguishing by race: LST white, LST non-white, PN white, and PN non-white. White students show the predictable pattern of increase across the three post-tests. Non-whites, however, look very different. Those in the LST group show a sharp decrease in use from time 1 to time 2, which leaves them with the lowest use level at the end of year three. Non-whites exposed to Project Northland, on the other hand, show a dramatic increase in use between the end of 7th grade and the end of 8th grade, with 14% reporting use at that point.

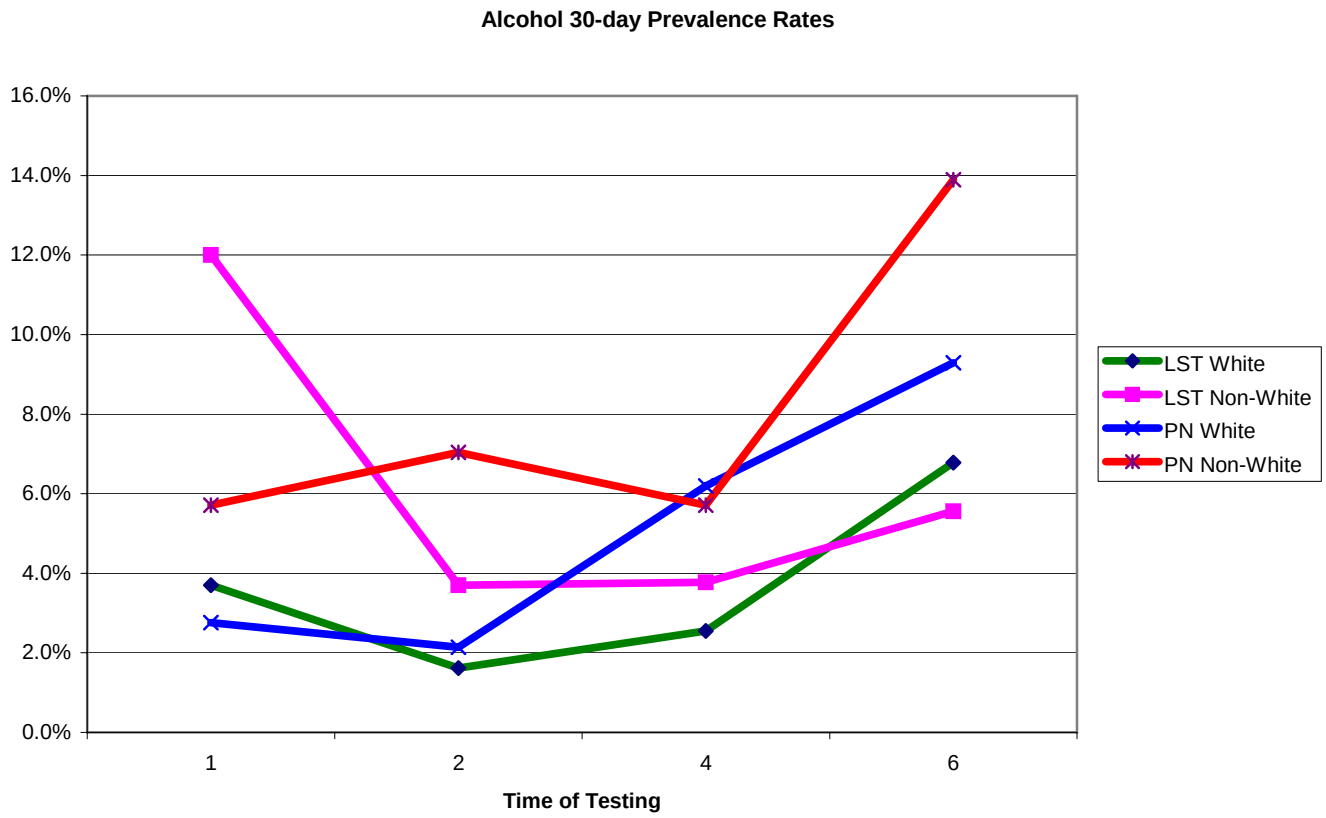

Figure 11. 30-day Alcohol Prevalence by Race

CONCLUSIONS

Did the programs reach the intended target population?

Both school-based programs reached relatively large numbers of students, with 669 in the 8th grade Project Northland treatment group and 657 in the 8th grade Life Skills Training treatment group, for a total of 1,326 students having received one of the prevention curricula at the time we conducted outcome analyses. In the matched sample for whom we have data tracked across all three years of the interventions, there were 536 in Project Northland and 485 in Life Skills Training, for a total of 1,021. Clearly we have enough statistical power to detect even modest effects for both of the classroom curricula.

For our comparison of treated 8th graders to comparable untreated eighth graders from the same schools, we checked the demographic similarity of these two groups and found that they were generally similar on most demographic variables we examined (gender, age, race/ethnicity, eligibility for subsidized lunch (an indication of socioeconomic status), family composition, and average grades). For example, both groups

were approximately 50% female, and ranged from 67.5% to 83.6% white, 55.3% to 79.7% ineligible for subsidized lunch, 71.5% to 79.7% in two-parent families, and 57.1% to 63.5% with grades mostly B or better. Looking at the two programs separately, there was a noticeable difference between treatment and comparison groups in the LST schools on race/ethnicity (higher percent white for the comparison group) and full-pay lunch (higher percent full-pay for the comparison group). One school in our sample was strikingly different from the others (lower percent white, and higher eligibility for subsidized lunch) and this led to an overall difference between the Life Skills Training sample (which included this school) and the Project Northland sample. The non-random way in which schools chose the curricula make any comparisons of the two classroom curricula suspect, particularly in the light of the noted demographic differences between the samples. For example, 75.8% of the Project Northland treatment group paid full price for their lunches, whereas only 55.3% of the Life Skills Training treatment group did so.

For the analyses in which we looked at changes from our initial 6th grade pre-test (Time 1) to our final 8th grade post-test (Time 6), we checked on the attrition in our matched sample. The participants for whom we could link Time 1 to Time 6 were demographically different from the students we assessed at Time 1 but were unable to match at Time 6. Consistent with the literature, attrition was non-random and the students lost from the study were lower in Percent white, full-pay lunch, two-parent family, and grades of B or better. This limits the generalizability of the findings for these analyses. Because of this, our 8th grade comparisons are more representative of the overall school populations.

For the Iowa Strengthening Families Program, we had 98 youth and 106 parents for whom we could analyze results matching pre-tests to post-tests for the first 7-session unit of the treatment. This represents only 7.4% of the 1,326 8th graders in our treatment sample for the classroom curricula. We found that the students in the ISFP program were more likely to be white and less likely to be in two-parent families than the youth treatment cohort as a whole. The relatively small sample in this program reduced our ability to detect statistically significant effects; furthermore, the subgroup of youth who participated was not entirely representative of the larger treatment cohort.

Were the programs delivered with fidelity?

All of the reports by our educators on their success in following the curriculum plans provided by program developers indicated relatively good fidelity. The highest fidelity was for the ISFP program (94.9%), followed by PN (88.4%). With the most sessions, LST was somewhat lower in fidelity at 80.1%, still respectable.

What were the properties of our measures of program outcomes?

We employed a number of measures in order to capture the effects of the programs on more proximal outcomes (intermediate objectives) likely to be directly affected by the programs, as well as the effects on the RFP-specified outcome objectives (30-day prevalence of drug use, initiation of drug use, and problem drinking). Some of the measures of intermediate effects were administered to all of our participants ("shared" measures) while others (designated "non-shared") were chosen to be specific to one of the programs (PN, or LST). For the drug use outcomes we focused primarily on alcohol because it is the most prevalent for this age group and is often viewed as a "gateway drug." In measuring intermediate effects for the ISFP program we chose five constructs and had both parents and youth report on them with slightly different items. The scales were drawn from a number of well-known sources and were uniformly acceptable in internal consistency reliability (Cronbach's Alpha) with our sample. In several cases the distributions of responses were non-normal and required transformations for our statistical analyses.

How did the programs do in the first year?

Effects on Year One objectives for changes on program-specific intermediate outcomes were disappointing for all three programs. One objective out of 14 was achieved for the classroom curricula: social skills (as reported by the youth) increased 19.5% from pre- to post-test for the LST group. None of the five objectives for ISFP was achieved. One likely problem is that the objectives were set too optimistically. Major research studies used in developing model programs are likely to produce much stronger effects than implementations in "real-world" contexts such as ours in this project. However, some of the trends were in the wrong direction (e.g. stress management, perceived peer use, and youth reports of parental expectations of non-use). In the light of the more promising results we will report from the 8th grade comparison group design, it may be valuable to reflect on the limited utility of single-year pre-to post-test comparisons as a means of demonstrating program effects. Adolescents are likely to be headed in the negative direction on many of the risk and protective factors most likely to be measured in substance abuse prevention evaluation, and even relatively powerful programs may not reverse these normal developmental trends – just slow them down.

Did the programs achieve intermediate objectives by eighth grade?

When we compared our 8th grade treatment group to untreated 8th graders from the same schools, we found that there was a highly significant difference for the seven "shared" outcome measures, indicating significant benefit from the programs. Six of the seven scales showed positive effects, with the largest differences for Favorable Attitudes Toward Drug Use and Drug Use Intentions (both lower for our treatment group). Only Family Attachment did not show a significant effect.

All four intermediate objectives specific to Project Northland showed modest but significant effects in the right direction. This included effects on parents (youth-perceived quality of parent communication and rule enforcement for ATOD use). For Life Skills Training, three of five intermediate outcome measures showed modest positive effects (higher drug refusal skills, reduced pro-drug attitudes, and lower perceived peer norms). There were no significant effects for the broader social functioning measures (social skills and stress management).

We also examined the changes from Time 1 to Time 6 for two intermediate outcomes we thought might show positive effects over time: Family Attachment and Parental Attitudes Toward Drug Use. We found that after some initial effects of the programs in the first year on parental disapproval of use, natural developmental changes appeared to dominate and perceived disapproval declined steadily. There were no program effects on attachment, which also declined over time, as we would expect for adolescents.

What about ISFP effects on intermediate outcomes? Parents indicated significant improvement over time on all five outcomes, but their offspring disagreed on four of the five. Only one intermediate outcome showed significant positive change for both parents and youth in ISFP: there was a significant increase in "Limit Setting and Monitoring." This is worth celebrating, as it is an important protective factor.

Did the programs achieve effects on substance use outcomes?

This question represents the "bottom line" for this project. We focused primarily on comparisons between treated and untreated eight graders to address it. For 30-day prevalence of alcohol use, probably the most widely chosen indicator for studies with this age group, both programs produced substantial effects (45% lower alcohol use than for the comparison group) that were highly significant and did not

differ between the two classroom curricula. We also used SALT data to confirm this finding, demonstrating that eighth graders in our intervention schools did show increasing reductions in alcohol use over the past three years while the rest of the schools in the state experienced a much more modest decline. For a second important outcome measure, initiation of alcohol use during the three years of the programs, there was a significant effect when both programs were combined but this was due to the substantial effect of Project Northland (42% lower initiation than the comparison group) and did not show up for the Life Skills Training intervention. We also looked at whether adding the Iowa Strengthening Families Program to the classroom curricula increased the effects on these outcomes, and did not find a significant added effect. However, there are important qualifications for this conclusion, including the small sample size and the weak self-report measure we had of participation in ISFP for these analyses.

Did the two youth-oriented programs affect demographic groups differently?

To check on how the programs did with youth from differing ethnic and economic backgrounds, as well as how the programs might have affected boys and girls differently, we conducted some analyses that examined effects on our two primary outcome measures, 30-day prevalence of alcohol use and initiation of alcohol use, for these subgroups. There were some interesting differences in program effects, and these also varied by outcome measure. To summarize briefly: (1) for 30-day prevalence, LST had greater positive effects than PN for white and full-pay (i.e. higher S.E.S.) students, while PN had greater positive effects on male students than LST did; (2) for initiation of use, PN had significantly more positive effects than LST for non-white students, lower S.E.S. students, and both boys and girls.

What final conclusions do we draw about program effectiveness?

It might help to think of "initiation" as referring to <u>any</u> experimentation with alcohol, while "30-day-prevalence" refers to more regular use over time. If reducing 30-day prevalence is the goal, <u>both</u> programs did very well and LST was especially effective for white and higher S.E.S. students; if the goal is decreasing initiation of use PN is a more promising choice, especially for non-white and lower S.E.S. students.